

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 5, 2003, 23:24:25 ; Search time 4104 Seconds
(without alignments)
6516.929 Million cell updates/sec

Title: US-09-914-168-2
Perfect score: 4/27
Sequence: 1 MSKPVLFANRSMFVALAAY.....TGKKEGNPKIKHFHFGTPTF 919

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPROO/US09914168/runat_28042003_151614_6787/app_query.fasta_1.1095
-DB=genembi -QFMT=fastcap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09914168 @CGN.1.1.2655 @runat.28042003_151614_6787 -NCPU=6 -ICPU=3
-NO.XLPYX -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vi:*
30: em_hg_hum:*
31: em_hg_inv:*
32: em_hg_other:*
33: em_hg_mus:*
34: em_hg_pln:*
35: em_hg_rod:*
36: em_hg_mam:*
37: em_hg_vrt:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4727	100.0	2760	6	AX034847	AX034847 Sequence
2	4727	100.0	89047	6	AX067459	AX067459 Sequence
3	4581	96.9	2670	6	AX034849	AX034849 Sequence
4	617	13.1	11158	1	AE004682	AE004682 Pseudomon
5	604.5	12.8	1374	6	AX078576	AX078576 Sequence
6	532	11.3	10302	1	AE006218	AE006218 Pasteurel
7	494	10.5	10831	1	AE013668	AE013668 Yersinia
8	494	10.5	216050	1	AJ414157	AJ414157 Yersinia
9	486.5	10.3	24892	1	AE008906	AE008906 Salmonell
10	486.5	10.3	249050	1	AL627283	AL627283 Salmonell
11	454	9.6	14390	1	AE004323	AE004323 Vibrio ch
12	436.5	9.2	1737	6	AX033468	AX033468 Sequence
13	436.5	9.2	11833	1	U32752	U32752 Haemophilus
14	434.5	9.2	243184	1	AP002568	AP002568 Escherich
15	433.5	9.2	10819	1	AE000493	AE000493 Escherich
16	433.5	9.2	11760	1	AE005654	AE005654 Escherich
17	433.5	9.2	338534	1	EC00963	U14003 Escherichia
18	427.5	9.0	1731	6	AX033470	AX033470 Sequence
19	375.5	7.9	11870	1	AE012531	AE012531 Xanthomon
20	359.5	7.6	10839	1	AE012071	AE012071 Xanthomon
21	353.5	7.5	20310	1	AY010120	AY010120 Xanthomon
22	347	7.3	10074	1	AE003957	AE003957 Xylella f
23	319	6.7	944	6	AX473246	AX473246 Sequence
24	268	5.7	201050	1	AL646064	AL646064 Ralstonia
25	265.5	5.6	2379	6	AX155443	AX155443 Sequence
26	265.5	5.6	2379	6	AX202493	AX202493 Sequence
27	262.5	5.6	10029	1	AE010493	AE010493 Fusobacte
28	259.5	5.5	2379	1	NG081959	U81959 Neisseria g
29	259	5.5	2394	6	AX155439	AX155439 Sequence
30	259	5.5	2394	6	AX155447	AX155447 Sequence
31	259	5.5	2394	6	AX202489	AX202489 Sequence
32	259	5.5	2394	6	AX202497	AX202497 Sequence
33	259	5.5	13538	1	AE002375	AE002375 Neisseria
34	259	5.5	110000	2	AC020884	Continuation (2 of
35	259	5.5	340806	1	NMA122491	AL162752 Neisseria
36	259	5.5	349980	6	AX043922	AX043922 Sequence
37	255.5	5.4	6444	1	AF461795	AF461795 Bartonell
38	253	5.4	2394	1	AF021245	AF021245 Neisseria
39	245.5	5.2	181233	2	AC010535	AC010535 Homo sapi
40	244	5.2	10029	1	AE012236	AE012236 Xanthomon
41	242.5	5.1	189050	1	AL646069	AL646069 Ralstonia
42	241.5	5.1	10242	1	AE011773	AE011773 Xanthomon
43	240	5.1	333800	1	SM591792	AL591792 Sinorhizo
44	238.5	5.0	6617	6	AX139467	AX139467 Sequence
45	236.5	5.0	10029	1	AE008175	AE008175 Agrobacte

RESULT 1

ALIGNMENTS

AX034847
LOCUS AX034847 2760 bp DNA linear PAT 15-NOV-2000
DEFINITION Sequence 1 from Patent WO0052042.
ACCESSION AX034847
VERSION AX034847.1 GI:11190801
KEYWORDS
SOURCE Moraxella catarrhalis.
ORGANISM Moraxella catarrhalis.
Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
Moraxella.
REFERENCE
AUTHORS Ruelle, J. L.
TITLE Immunogenic compounds
JOURNAL Patent: WO 0052042-A 1 08-SEP-2000;
RUELLE JEAN LOUIS (BE) ; SMITHKLINE BEECHAM BIOLOG (BE)
FEATURES
source location/Qualifiers
1..2760
/organism="Moraxella catarrhalis"
/db_xref="taxon:480"
BASE COUNT 775 a 345 c 654 g 786 t
ORIGIN
Alignment Scores:
Pred. No.: 5,68e-289 length: 2760
Score: 4727.00 matches: 919
Percent Similarity: 100.008 conservative: 0
Best Local Similarity: 100.008 mismatches: 0
Query Match: 100.008 indels: 0
DB: 6 gaps: 0
US-09-914-168-2 (1-919) x AX034847 (1-2760)
OY 1 MetSerLysProValLeuPheAlaAsnArgSerPheMetProValAlaLeuAlaIleAtyr 20
DB 1 ATGCAAGCCCGCTTTGCTTGTGCAAAATCGAGTTTATGCTGCGATGCGCGCTAT 60
OY 21 LeuProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsn 40
DB 61 TTGCGCTTGATGACATCGCAAGCATTTGGCCACACAAATTAACCTCGCAACATCATCAT 120
OY 41 HisValProAlaHisAspThrAlaIleAsnGlnAlaLysAlaGlnAsnProProValLeu 60
DB 121 CATGTACCCCGCTCATGACACCGCATCAATCAAGCAAGAGGCAATCCCGCTTTTG 180
OY 61 LeuThrProGlnGlnIleGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGln 80
DB 181 CTAACACCTGAGCAGATACAAAGCAGCGCTTAATGCTGCTGACATGGAATGCTAACGCCCAA 240
OY 81 SerGlnAlaLeuAspValIleAsnPheAspAspGlnSerProIleSerArgIleGlyGlu 100
DB 241 TCACAAAGCTTTGATGTTGCTCAATTTGATGATCAATGCGGATATCTCGATCGGTAG 300
OY 101 GlnSerProProLeuGlyLeuAspMetSerValIleGlnGlnIleThrThrProLeuSerLeu 120
DB 301 CAATCACCCCGCTTTGGTTGGATGATGTCGCTCATCGAAGAAACACACCGCTAAAGCTTG 360
OY 121 GlnGluLeuPheAlaGlnGlnIleSerThrGlnMetGlyIleAsnProAsnAspTyrIlePro 140
DB 361 GAGCAATTAATTTGCTCAAGATCTACTGAGATGGAATCAATCCAAATGATTATATTTCCA 420
OY 141 GlnTyrGlnGlnGlnGlnIleProAsnSerGlnValValValProProThrLeuGlnProGln 160
DB 421 GAATATATCAAGCGCAGCAACCTTAATAGTAGGTGTTGACACCGACCTTAAGAACTGAA 480
OY 161 LysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysVal 180
DB 481 AAACACGAGTTTGAATCAAGCGCTTTATGACAGCGCTATTATGATGGTCAATAAGGTG 540
OY 181 ProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGlnThrSerAlaIleGly 200
DB 541 CTTAAGCTTAAGGCAAAATTTATCAATCATCGCAATCAGCGAAGACAGTGCATTTGGG 600
OY 201 SerSerIleGlnLysThrGlnLysProTyrAlaAsnIleLysAlaIleLeuGlnAspIleThr 220

DB 601 TCATCGCATCAAAAAACAGACCTTATGCAATATCAAAAGCAGCAGCTGAAGACATCAC 660
OY 221 GlnGlnSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuVal 240
DB 661 CAAAGTACGAGTATGATGATTTAAATGGCTTATCCACCGCTTAAGGCAAAACGCTTTGGTG 720
OY 241 AlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGly 260
DB 721 GCACGCGCTGCTGTCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 780
OY 261 GlnValAspValIleIleHisAspLeuGlyLysProValTyrIleAspTyrArgAlaVal 280
DB 781 GAGGTGATGATCATCATCCATGATGATGATGATGATGATGATGATGATGATGATGATG 840
OY 281 GlnValArgGlyGlnGlnAlaAspAspLysAlaPheThrThrThrValAlaAspGluValPro 300
DB 841 GAGGTACAGGTGAAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
OY 301 LeuLeuIleGlyAspValPheHisHisGlyLysTyrGlnThrLysLysAsnLeuIleGlu 320
DB 901 TTGCTGATCGCGCATGCTGCTTATCATCATGCAAGTACGAACCAAAATCTCATCGAA 960
OY 321 AsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgTyrPheAspAspArgSerValAspVal 340
DB 961 AATGCCACGTGCTGAACATGATTTGATGATGATGATGATGATGATGATGATGATGATGATG 1020
OY 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlyTyrArg 360
DB 1021 ATTTTCCAGATTAATACCGCTGATGCTGATGATGATGATGATGATGATGATGATGATG 1080
OY 361 PheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
DB 1081 TTGTATGAGGTGATATTTTATTTTACATGATGATCTTAAACCAATGATGCAACAGCATCCA 1140
OY 381 AspLysLeuProValLysArgGluLeuLeuGlnGlnLeuLeuThrValAsnMetGlyGlu 400
DB 1141 GATTAAGCTGCCAGTTAAACGAAATTTACTTGAGCGTTACTACCGCTTAACATGAGAGAG 1200
OY 401 AlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
DB 1201 GCTTACATTTTACAGGCGGTGCTGACTTTCATATGATTTGATTTGCCACAGGATTTT 1260
OY 421 AsnMetValAsnThrGlnIleValPheProGlnArgGlnGlnIleGlnAsnAspGlnVal 440
DB 1261 AATATGGAATACCGAATGATGCTTTCCAGAGGCTGAACGATGCAAAAGCACCAATG 1320
OY 441 SerPheGlnGlnSerSerSerSerArgThrGlnProAlaGlnValAspGlnSerThrLeu 460
DB 1321 AGCTTTGAGCGTCTTCAAGTAGCCGTAAGTACAGCAGCAAGTGAAGAGAAAGCACACTT 1380
OY 461 GluProValIleGlnThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIle 480
DB 1381 GAACCTGCTATGAACCGTTGACGCTAAGCGATGGAGATTAATAGATTTTCCGCCATC 1440
OY 481 GluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArg 500
DB 1441 GAATTTAGTGCATCTAATCTGATTCAGACCAAGCAAGCTTAATTTGGCGCGCCAAAGCTCGC 1500
OY 501 HisLeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
DB 1501 CATTTATATGACATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
OY 521 ArgSerIleLeuGlyArgIleSerAspAlaValAlaSerAlaValAlaArgAlaIleLeuPro 540
DB 1561 CGCTCTATTTTGGGAGAGATGAGGATGCCGTATCTGCGTTGCACGCTATTTTACT 1620
OY 541 AspGlnSerGluAsnGluValIleAspLeuProGlnArgThrAlaLeuAlaAsnArgLys 560
DB 1621 GATCAATCTGAATAATGAGGTAAATGATTTGCCAGAGCTACCGCATTTGGCTAAATCGCAAG 1680
OY 561 ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValAlaPheValAlaSerAsp 580

Dh 1681 ACCCTGCTGATGCTATCAAGTAAAAAGTCCGCTATATCTCTTGTGGCGAGTGAT 1740
Qy 581 LysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuVal 600
Dh 1741 AATCCACGAGATGTCGCAATTTGTTGGCTGGGATCGGACACAGTACCGGCTAGTC 1800
Qy 601 ThrLysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaLeuLeuArg 620
Dh 1801 ACAAAATTTGACATTAATTTGATTCGTATCGTATGCTATCAAGACGCGCTAGCTAAGA 1860
Qy 621 LeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeu 640
Dh 1861 CTGTCGAGGATAAAAAAGGGGTCAAGTTATATGCCACCAACCCCTTAGCCACCTCTA 1920
Qy 641 AsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGlnValPheGlyHisSerThrAsn 660
Dh 1921 AATGATCACTAGACAGCAACTTTGGGTATCAACAAGAAATTTTGGTCACTTACCAAT 1980
Qy 661 GlyPheAspLeuSerThrArgThrLeuGlnHisGlnIleSerArgSerIleIleGlnAsn 680
Dh 1981 GGTTCCTGATTTATCCACACGACACCTAGAGCATGAGATTAGCCGAGTATTCACAAAT 2040
Qy 681 GlyGlyTrpAsnArgThrThrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAla 700
Dh 2041 GGTGGCTGAATCGTACTTATTCATTCGTTATCGTTCGTGATTAACCTTAAACCCACAGCA 2100
Qy 701 ProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysProSerGlnIu 720
Dh 2101 CCCCCGTAACATGCGAGATTTACCAGTGGATTTTGTCATAGTAAAGCCACCAAGG 2160
Qy 721 AlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMet 740
Dh 2161 GCGTATTATGGCAGCGTGTCTGTCATAAACGTTGCGATTAATTCGATTCGATG 2220
Qy 741 ArgGlyTyrArgGlnArgTyrSerLeuGlnValGlySerSerGlyLeuValSerAspAla 760
Dh 2221 CCGGCTATCGTCACGATATTCCTTAAAGTGGCTCAAGCGGTTTGGTATCGATGCT 2280
Qy 761 AsnMetAlaIleAlaArgAlaGlyLysSerGlyValTyrSerPheGlyAspAsnAlaTyr 780
Dh 2281 AATATGGCATATTCGTCAGCTGGTATAGTGGCGGTATAGTTTGGGATATATGCTTAT 2340
Qy 781 GlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Dh 2341 GGCACCAATCGTCCCATCAGATGACTGGTGGCATCAAGCAGATACATTGTCGCGAT 2400
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyLysAspGlnSerIleArg 820
Dh 2401 AATTTTAATCATGTCCCATATCGTTTGGCTTTTGGCTGGTGGCGACCAAGTATGCT 2460
Qy 821 GlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGln 840
Dh 2461 GGAATGACATGACAGACATTTATACCTATATCAGATTAAGGTATATCTGACAGCGGTCA 2520
Qy 841 ValLeuAlaValaGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAla 860
Dh 2521 GATATGGCGGTGGTACAGCTGAATATATATGAATTATGAAGATTGCGTTGGCG 2580
Qy 861 ValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGly 880
Dh 2581 GTTTTGGTATATGATATGATCTTATGATTAAGGCTTTTACTAAATGATACCAAAATGCT 2640
Qy 881 AlaGlyValaGlyValaArgTyrAlaSerProValaGlyGlnValaArgValaAspValaThr 900
Dh 2641 GCAGGTGTCGCTGCTCGCTGGGACATCACCCTGCGCTCAAGTTGCTGTTGATGTCGCACT 2700
Qy 901 GlyValLysGluGluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
Dh 2701 GGTGTCAAAGAAAGGAGGCAATCCATTAAAGTCATTTTATATGACACACCATTT 2757
RESULT 2
AX067459
LOC05 AX067459 89047 bp DNA linear PAT 24-JAN-2001

DEFINITION Sequence 34 from Patent WO0078968.
ACCESSION AX067459
VERSION AX067459.1 GI:12545079
KEYWORDS
SOURCE Moraxella catarrhalis.
ORGANISM Moraxella catarrhalis
Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
Moraxella.
REFERENCE 1 (bases 1 to 89047)
AUTHORS Lagace,R.F., Patterson,C. and Berg,K.L.
TITLE Nucleotide sequences of moraxella catarrhalis genome
JOURNAL Patent: WO 0078968-A 34 28-DEC-2000;
Incyte Genomics, Inc. (US)
FEATURES
source location/Qualifiers
1..89047
/organism="Moraxella catarrhalis"
/db_xref="taxon:480"
unsure
BASE COUNT 26501 a 17338 c 19060 g 26147 t
ORIGIN
Alignment Scores:
Pred. No.: 5,56e-287 Length: 89047
Score: 4727.00 Matches: 919
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-914-168-2 (1-919) x AX067459 (1-89047)
Qy 1 MetSerLysProValLeuPheAlaAsnArgSerPheMetProValAlaLeuAlaTyr 20
Dh 80397 ATGTCAAAACCCGCTTTGTTGTTCCAAATCGCACTTTATGCCCTGCGCATTTAT 80456
Qy 21 LeuProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsn 40
Dh 80457 TTGCCCTTGATGATGATCGCAACATTCGGACCAACAAATTAACCTGCAACATCATAT 80516
Qy 41 HisValProAlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyLysAsnProProValLeu 60
Dh 80517 CATGTACCCGCTCATGACACCCCATCAATCAAGCAAGGAGCAATCCGCTGTTTGG 80576
Qy 61 LeuThrProGluGlnIleGlnAlaArgLeuAsnAlaAlaGlyLysAsnAlaLysProGln 80
Dh 80577 CTAACACCTGACGACATACAGACAGCCCTTAATGCTGCTGGAGCTAAATGCTAAGCCCA 80636
Qy 81 SerGlnAlaLeuAspValaValaAsnPheAspGlnSerProIleSerArgIleGlyGln 100
Dh 80637 TCACAAGCTTTGATGTTGTCAATTTTGTATGATCAATCGCCGATATTCGTTGAG 80696
Qy 101 GlnSerProProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeu 120
Dh 80697 CAATCACCCCTTTGGTGGTATGTCGGTCATCGAAGAAACACACCCGCTAAGCTTG 80756
Qy 121 GlnGluLeuPheAlaGlnGlnSerThrGlnMetGlyIleAsnProAsnAspTyrIlePro 140
Dh 80757 GAGGATTTATTTGCTCAAGAAATCTACTGAGATGGGAATCAATCAAAATGATTTATTTCA 80816
Qy 141 GlnTyrGlnGlyGluGlnProAsnSerGluValaValaProProThrLeuGluProGlu 160
Dh 80817 GAATATCAAGGCGACCAACTTAATAGTGAAGGTGTGATCCACGCACTTAACAACTGAA 80876
Qy 161 LysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValaLysLysVal 180
Dh 80877 AAACCAAGGTTGATCAACGCTTTATATGCACCCCAATTAATGATGCTCATTAAGGTC 80936
Qy 181 ProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGly 200
Dh 80937 CTTAGGCTTAAAGCAAAATTTTATATCATTCGCATACGAGGCAACCAAGTGGCTTTGG 80996
Qy 201 SerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThr 220

Db	82077	ACCCTGCTGATGCTGTATCAAAAGTAAAGTGGCCGATATGCTCTTGTGGCCAGTAT	82136
Qy	561	LysProArgAspGlyGlnIleGlyLeuGlyTyrPglSerAspThrGlyThrArgLeuVal	600
Db	82137	AAACCCAGCAGATGGTCAAAATTGGTTGGGCTGGGATGCGACACAGGTAAACCCGGCTAGTC	82196
Qy	601	ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnIleGlyValGluLeuArg	620
Db	82197	ACAAATTTGACATAAATTGTGATTATGCTATGGCTATCAAGCAGGGCGTGGACTAAGA	82256
Qy	621	LeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeu	640
Db	82257	CTGTCTGAGGATTAATAAAAGGGGTCAAGTTATATGCCACCAACCGCTTAGCCACCTCTA	82316
Qy	641	AsnAspGlnLeuAlaGlyAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsn	660
Db	82317	AATGTCACGCTAAGAGCAACTTTGGGTTATCAACAAGAAAGTTTGGTCACCTTATCCAT	82376
Qy	661	GlyPheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsn	680
Db	82377	GGTTTTGATTTATCCACACGCACCTTACAGCATGTGATTACCCGACGATTTATCCAAAT	82436
Qy	681	GlyGlyTyrPAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAla	700
Db	82437	GGTGCTGCGAATCGACTACTTATTCATTCGCTATCCGTTACGTTGAAGCTTAAACCAACA	82496
Qy	701	ProProGluPrrTrpGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGlu	720
Db	82497	CCCCGTGAACATGGCAGGATTTTACCAGTGGATTTTGTTCATATGGTAACCCAAACCAAG	82556
Qy	721	AlaLeuLeuAlaGlyValAlaValAlaHisLysThrValAlaAspAsnLeuValAsnProMet	740
Db	82557	GCGTTATTGGCAGAGGTGTGCTGTGCATTAACACGGTCCAGATTAATTTGGTTAAATCCGATG	82616
Qy	741	ArgGlyTyrArgGlnArgTyrSerLeuGlnValAlaGlySerSerGlyLeuValSerAspAla	760
Db	82617	CGTGCTATTCGTACACCGATATTCCTTTACAGAGTTGGCTCAACGGTTTGGTATCCGATCT	82676
Qy	761	AsnMetAlaIleAlaArgAlaGlyLleSerGlyValTyrSerPheGlyAspAsnAlaTyr	780
Db	82677	AATATGGCTATTGGCCGAGCGTATGTATGGCGGTATAGTTTGGGGATAATCCCTAT	82736
Qy	781	GlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleIleTrpSerAsp	800
Db	82737	GGCACACATCGGCCCATCAGTATGCTGTGGCATACAGCAGGATACATTTGGCGGAT	82796
Qy	801	AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg	820
Db	82797	AATTTATTCATGTGCCATATCGTATCGTTTGGCTTGGTGGGACCAAGTATTCGT	82856
Qy	821	GlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGln	840
Db	82857	GGATATGCACATGACAGATTATACACTTATATCAGATAAGGGTTATCTACAGCGCGTAA	82916
Qy	841	ValLeuAlaValAlaGlyThrAlaGlyTyrTrpAsnTyrGluPheMetLysAspLeuArgLeuAla	860
Db	82917	GTAATTGGCGGTTGGTACACGCTGAATATATATATGATATATGAAAGATTTGGCTTTGGCG	82976
Qy	861	ValPheGlyAspLleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysLleGly	880
Db	82977	GTTTTTGGTATATGCTAATGCTTAATGCTTAATGAAGGCTTTACTAATGATNCCAAATTTGCT	83036
Qy	881	AlaGlyValAlaGlyValArgTrrAlaSerProValGlyGlnValArgValAspValAlaThr	900
Db	83037	GCAGGTGTCGGGTTCGCTGGGCAATCCACTGCGGTCAAGTTCCGTGTGATGTGGCAACT	83096
Qy	901	GlyValLysGlnGlnGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe	919
Db	83097	GGTGTCAAAAGAAAGAGCGCAATCCATTAAAGTCGATTTTTTTATATGGCACACACATT	83153

Tue May 6 09:37:16 2003

us-09-914-168-2.rge

age 5

Sequence 3 from Patent WO0052042.
ACCESSION AX034849 GI:11190802
DEFINITION Moraxella catarrhalis.
KEYWORDS Moraxella catarrhalis.
SOURCE Moraxella: Proteobacteria; gamma subdivision; Moraxellaceae;
ORGANISM Bacteria.
Moraxella.
1 (bases 1 to 2670)
REFERENCE 1 (bases 1 to 2670)
AUTHORS Ruelle, J. L.
TITLE Immunogenic compounds
JOURNAL Patent: WO 0052042-A 3 08-SEP-2000;
RUELLE, JEAN LOUIS (BE) ; SMITHKLINE BEECHAM BIOLOG (BE)
FEATURES
source location/Qualifiers
1.2670
/organism="Moraxella catarrhalis"
/db-xref="taxon:480"
BASE COUNT 757 a 527 c 632 g 754 t
ORIGIN
Alignment Scores: 8,72e-280 Length: 2670
Pred. No.: 4581.00 Matches: 888
Score: 100.00% Conservative: 1
Percent Similarity: 99.89% Mismatches: 0
Best Local Similarity: 96.91% Indels: 0
Query Match: Gaps: 6
DB: 6
US-09-914-168-2 (1-919) x AX034849 (1-2670)
OY 31 GlnGlnAsnProAlaAsnTlleAsnHisValProAlaHisAspThrAlaIleAsn 50
Db 1 CAACAAATTAACCTGCAAAATCATCATCATCATCATCATCATCATCATCATCATCAT 60
OY 51 GlnAlaValAlaGlnAsnProValLeuLeuThrProGlnInleGlnAlaArgLeu 70
Db 61 CAGCAAGAGCGAGCAATCCCTGTTTGTCTACACCTGAGAGAGAGAGAGAGAGAGAG 120
OY 71 AsnAlaAlaGlnLeuAsnAlaValProGlnSerGlnAlaLeuAspValAlaAsp 180
Db 121 AATCTCTGCTGACTGAATGCTTAAGCCCATCAACAGCTTGATGCTCAATTTTGA 110
OY 91 AspGlnSerProIleSerArgIleGlyGlnInleuProIleuGlnLeuAspMetSer 240
Db 181 GATCAATCGCGATATCTCTATCGGTGAGCAATACCCCTTGCTGATATCTCG 130
OY 111 ValIleGlnGlnThrProLeuSerLeuGlnGlnGlnGlnGlnGlnGlnGlnGln 300
Db 241 GTATCGAAGAAACACACCCCTGAGCTGAGAGATTTTGTCTCAAGATCTACTGAG 150
OY 131 MetGlyIleAsnProAspArgIleProGlnGlnGlnGlnGlnGlnGlnGlnGln 360
Db 301 ATGGGAATCAATCAAAATGATTTATTCAGAAATATCAAGCGCAAGCACTTAATG 170
OY 151 ValValValProProThrLeuGlnProGlnGlnGlnGlnGlnGlnGlnGlnGln 420
Db 361 GTGGTTGTACACGACATTTGAGACCTGAAACCGAGTTTGATCAAGCTCTTATGCA 190
OY 171 ArgLeuPheAsnAspGlyValAlaAsnValProArgLeuValAlaValSpheryGlnSer 480
Db 421 CCCATTATTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 210
OY 191 SerGlnSerGlyGlnThrSerAlaIleGlySerSerHisGlnGlnGlnGlnGln 540
Db 481 TCCCAATCGAGCAACAGAGCGATGGTCAATGCAATCAAAACAGAGCGCTTATG 230
OY 211 AsnIleValAlaLeuGlnAspIleThrGlnGlnGlnGlnGlnGlnGlnGlnGln 600
Db 541 AATATCAAGAGCACTGATGAGACATCCCAAGAGTCAAGCGATTTGATGAGCTCT 250
OY 231 IleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIle 250
Db 601 ATCCAGCGCTAAGCAAACTGCTTGTGCGAGCGCGCTGCTGCTGCTGCTGCTGCT 660

OY 251 AspleuSerIleIleArgAsnSerIleGlyGlnValAspValIleIleHisAspLeuGly 270
Db 661 GATTATCAATCAATTAAGAAATAGCATCGAGAGAGTGCATCATCATCATCATCATCAT 720
OY 271 GlnProValTyrIleAspTyrArgAlaValGlnValArgGlyGlnValAlaAspAsp 310
Db 721 GAACCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 310
OY 291 AlaPheThrThrValAlaAspGlnValProLeuLeuIleGlyAspValPheHisGly 840
Db 781 GCATTACTACCGGCGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 330
OY 311 LysTyrGlnThrLysLysAsnLeuIleGlnAsnAlaSerAlaGlnHisGlyTyrPheAsp 900
Db 841 AAGTACGAACCAAAATAATCTCATCGAAATCCAGTCCGATCAACATGATGATTTGAT 350
OY 331 GlyArgThrAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSer 960
Db 901 GGCGGTGGCTGATCGCTCGATGATGATGATGATGATGATGATGATGATGATGATG 370
OY 351 LeuIleTyrAspThrGlyThrGlnTyrArgPheAspGlnValAlaPhePheThrIleAsp 1020
Db 961 TTAATTAATGATACAGTACGAGATGATGATGATGATGATGATGATGATGATGATG 1080
OY 371 ProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGlnLeu 410
Db 1021 CCTAAACCAATCAATTAAGCAACCAATTAAGCAACCAATTAAGCAACCAATTAAG 1140
OY 391 GlnGlnLeuLeuThrValAlaAsnMetGlyGlnAlaTyrAsnLeuGlnAlaValArgAlaLeu 430
Db 1081 GAGCAGTACTCCGCTTAACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
OY 411 SerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGlnIleValPhePro 450
Db 1141 TCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
OY 431 GlnArgGlnGlnIleGlnAsnAspGlnValSerPheGlnGlnSerSerSerArgThr 470
Db 1201 GAGCGTCAACATCAAAAGCAACCAATGAGCTTTGAGCAGTCTCAATACCGCTACT 1320
OY 451 GlnProAlaGlnValAspGlnSerThrLeuGlnProValIleGlnThrValGlnLeuThr 490
Db 1261 GAAACCAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
OY 471 AspGlyIleLeuMetLysPheSerProIleGlnPheSerLysSerAsnLeuIleGlnAsp 510
Db 1321 GATGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
OY 491 LysLeuAsnLeuValAlaAlaValAlaValAlaValAlaValAlaValAlaVal 530
Db 1381 AAGCTAAATTTGGTGGCTGCCAAGGCTGCCATTTTATGATGATGATGATGATG 1500
OY 511 LeuAlaIleAsnHisAspAspGlyValAlaAsnArgSerIleLeuGlnArgIleSerAspAla 550
Db 1441 CTGGCAATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
OY 531 ValSerAlaValAlaAspAlaIleLeuProAspGlnSerGlnAsnGlnValAlaIleAspLeu 570
Db 1501 GTATCTGCGCTGACGAGTGTATTTTACCTGATGATGATGATGATGATGATGATG 1620
OY 551 ProGlnArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysLys 590
Db 1561 CCCGAGCTACCGATTTGGCTATGCAAGACCCCTGATGATGATGATGATGATGATG 1680
OY 571 ValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGly 610
Db 1621 GTGGCGCTATGCTCTTGTGCGAGATGATGATGATGATGATGATGATGATGATG 1740
OY 591 TrpGlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArg 650
Db 1681 TGGGATCGGACACAGCTACCCGCTAGTCAAAATTTGATGATGATGATGATGATG 1740


```

gene
CDS
    /translation="MLTQGLFRPLAGLCLSTPAFAESQLSVRYTPANALANIE
    AASGLASSLDLFFKRF"
    complement(4508..6247)
    /gene="PA2543"
    /complement(4508..6247)
    /gene="PA2543"
    /codon_start=1
    /transl_table=1
    /product="conserved hypothetical protein"
    /protein_id="AAG05931.1"
    /db_xref="GI:9948601"
    /translation="MLTQGLFRPLAGLCLSTPAFAESQLSVRYTPANALANIE
    AAYGLSEDEALQFRRNAEAQAEKAAQALGFQAOIDSEVDGKPKLTLVYPG
    EPVRLQVNIQVLEAASLESFRLPSGOLKPGAKLNGVYEDAKRLIONASRGGF
    OGFRSTQRLSDPAGIADIDLVYDSGORYFGKVSFDGSIIEELLIRMYPRKAG
    PYSELIAELNONQSSGPEGVDAAPTOAODGARQALPVAVREARKPRMGV
    LGRSTDGARARENMTWYNAEBSHGFEISAPRONGAMTEIPLDPLTLRF
    TSGVDFDLVDTESKLLTLGGEMHSKRPDQWVSLNMMREYKLGDDSLFSLFP
    GIGVSLKETDKVDPSHGRLQFNVKAGEGLADADVLHVDAMAKGTFPAGHRL
    GRLQVGIATNDYKSIPLSLRFAGGQSVRGYDRTLSPENSQDKIGGRYMIAGV
    EYQPLAEWRMLATFVDQGNFNSLDPSIKTGVCYGVWVSPVGLRLDLAHLDD
    GGRFLHSMGPEL"
    complement(6381..7094)
    /gene="PA2544"
    complement(6381..7094)
    /gene="PA2544"
    /codon_start=1
    /transl_table=1
    /product="hypothetical protein"
    /protein_id="AAG05932.1"
    /db_xref="GI:9948602"
    /translation="MARPSARPIAVAKRACACAGROGMRSPQVTVDRLLDGYR
    EVSLTFYAHRYHEPTFAYLMDQAGDQDQVACVREVRHLEELPAIGLFDRL
    VAAALITPPORLIDITESWSMRMLFSLVGLAARRYLQVRAVQACLPVGHQDPL
    LGVPEYQGRQIGBQLLAQVORMCEDSASGVLVDIGNARYLDFYQRHGRIGEVT
    LGAVREHVLHPGETPAT"
    7239..8051
    /gene="xtna"
    /note="PA2545"
    7239..8051
    CDS
    /gene="xtna"
    /codon_start=1
    /transl_table=1
    /product="exodeoxyribonuclease III"
    /protein_id="AAG05933.1"
    /db_xref="GI:9948603"
    /translation="MKIVSFNINGLRARPHQALAIRHOPDVIGLOETKADQFRO
    AEIETALGYHYHSGKHGYVALISROAPLELHGFPSDGESEGRFTWGFSDAQN
    PIVMNGYFPQGENREHPTKFPKQRFYADQSLNENFKPEQALVVMGDINISPEC
    DIGIENNRKRLTKGCSFLPEEREMLATLKNMGLVDSFALNPEVNDRESMFDRS
    RGFEDNPKRGIRIVILASRVLOGRFKDAGIDYDLRGMKPKSDHAPIMLELA"
    complement(8161..8592)
    /gene="PA2546"
    complement(8161..8592)
    /gene="PA2546"
    /codon_start=1
    /transl_table=1
    /product="probable ring-cleaving dioxygenase"
    /protein_id="AAG05934.1"
    /db_xref="GI:9948604"
    /translation="MTMQPFVERLDHIIVLRVDLERSIAFYRDVLGCVVKKMDLE
    LAHLRLGTSMILVAASGALGREGAARGGEGRWDFHCVAIEFFDEALVAHLEFG
    LVPQVAEIRFEGAGDGMSTLYFPDPCNRVELKPRVDSPA"
    8683..9600
    /gene="PA2547"
    8683..9600
    /gene="PA2547"
    /codon_start=1
    /transl_table=1
    /product="probable transcriptional regulator"
    /protein_id="AAG05935.1"
    /db_xref="GI:9948605"
    /translation="MNPVDHLTALKVRAVVAANGFAAARONNISPAVSKVAEL
```

```

EHLKRVLRINRTSRMSLTEAGEVYRORLERLDDLEAADAATSMOOGPSGLRVSA
PLTLATCTPAIPAFIORYPELRLLEILLQGRDOLIAEGDILRGSDRVADSGIVA
RPLVLELVICAAAPAYISORHQPRLPELRLREHRECIRESLSHADRWTRKRECIAY
IAGRIYSSSLAVRDALLAGFGLSLIPRLVYQAEIABGLVELLDWAKDEATLHAVY
PSROLACKTRVLELDTETMAOGHDSPTL"
9660..11054
/gene="PA2548"
9660..11054
/gene="PA2548"
/codon_start=1
/transl_table=1
/product="hypothetical protein"
/protein_id="AAG05936.1"
/db_xref="GI:9948606"
/translation="MPAPRHQRASDRKLNRSITSLVFLGFLCAFPMSVPAALPPI
TDDPALIQGSNTGTALPLVGLGELDEEGLDTRTPEREHEHLTKSRDGRIL
VRVELSANGSTFTFALLERIAELAAASRPVDSVRVLAIFGLRSPYAOQVLAIDG
LAILHPNPPLASIDTAQMAAVFAGEVHTWELGSGGAIPLVARDQSGTYDFKEL
VLASHGSLASDARFESNDRLAAVASDPAHIGVGIASVGRARALAIDGESLPM
VOESLATEDYELSRLEFLYLKPGKKNMAQALVRFAGPGQAIWTRSGVAKIOA
VQIAPRQMPAEYRIKLAEOARLITRPFQSGSALIDMKALLDQRLIDYLRQNRKLO
ERTLVGFGDKPSDKERSVLLSRLPAAVRELRANGVSFRELIGFDEELVASNDQE
LGRQNRREVEWVY"
BASE COUNT      1804 a      3911 c      3650 g      1793 t
ORIGIN

Alignment Scores:
Pred. No.:      4,69e-29      Length:      11158
Score:          617.00      Matches:      223
Percent Similarity: 37.64%      Conservative: 138
Best Local Similarity: 23.25%      Mismatches:   298
Query Match:    13.05%      Indels:      302
DB:              1      Gaps:          27

US-09-914-168-2 (1-919) x AE004682 (1-11158)
QY      35  ProAlaAsnIleIleAsnHisValProAlaHisAspThrAlaIleAsnGlnAlaLysAla 54
        |||||:::  ::  |||  |||  |||:::
Db      6723 CCTGGCATGCTGCTCAGCGTGGCGTGGCCGCGACCGCGGCTA-----CCT 6676
QY      55  GlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaArgLeuAsnAlaLacly 74
        ||  |||||
Db      6675 GGATTACCAACG-----GGCGGTCCAGGC 6652
QY      75  LeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspGlnSerPro 94
        |||
Db      6651 CTG-----CCT 6646
QY      95  IleSerArg-IleGlyGluGlnSerProProLeuGlyLeuAspMetSerVal----- 111
        ||  |||  |||  |||||:::
Db      6645 GCCGGTCCGCCCCCACCAACCAATTGCTCGGCGCTGCATCCCGAGTACCAGGCGCAG 6586
QY      112 ----IleGluGluThrThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrSl 130
        |||  |||  :::::  |||:::  |||
Db      6585 GCAGATCCGCGCAGCAGTTCGGCGCGCTGCAGCGCTGGTCCGCGAGAGACTCCGCTTC 6526
QY      130  uMetClyIle-----AsnProAsnAspTyrIleProGluTyrGln----- 143
        |||:::
Db      6525 GCAGGGGCTGCTTGATACCGGCAATGCGCGCTACCTGAGCTTACACGCGGCGACGG 6466
QY      143 ----- 143
Db      6465 CTACCGCGAGATCGCGCAGAGTGCAGCTGGAGCGGTCGCCGACACGTCCTCTGCATCC 6406
QY      144  -GlyGluGlnProAsnSerGluValValValProProThrLeuGluProGluLysProGl 163
        |||||  |||  :::::  |||||  |||||
Db      6405 TGGCGAGACGCCGCGC-GGCGACCTGAGCTCCCGCGACGCCGA--GAAGAACCATG 6350
QY      163  YLeuIleLysArgLeu-----TyrAl 170
        |||  ::|||  |||
Db      6349 CCTGTG-CAGGCGCTTGCGCTCTAAACCCCAATGTCAGAGAGGTGTGCGCGCGCTCCG 6291
QY      170  aArgLeuPheAsnAspGlyValAsn-----LysValProArgLeuLysAlaLys 186
```

```

Db 6290 GCAATCTGCACGCGCCACCCGCTGCGATCATCCGCCATGAGACTGACTCAAGG 6231
OY 186 sPhetYrClInserSerGInserGly-----GlnThrSerAlaIleGlySerSe 202
Db 6230 ATTGCTCCGGGTTCCCGCTGGCGGCGTCTGCTGCTGAGTACGGCGCTTCCGGGAGAG 6171
OY 202 rHis-----GlnYsThrGluProTyraAlaAsnIleLysAlaIalaLeuGluAsp11 219
Db 6170 CCACCTTACCGCTGGCGGTCACCCCGCAATGCGCCCTGAAAGCCATATCGAGGCTTA 6111
OY 219 eThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLe 239
Db 6110 TCTGCGTAGCCTCGCGAGCGGACGAAAGCGCTTGCAAGGTTTCCGCCCAATGCGCGA 6051
OY 239 uVal-----AlaAlaArgAlaValaGlyTyrrTyrrAspIleAspLeuSerIle11 255
Db 6050 AGCCGAGGCGGAAAAGCGCGCCAGCGCTCGGCTACTTCCAG----- 6008
OY 255 eArgAsnSerIleGlyGluValaAspValIleIleHisAsp----- 268
Db 6007 -----GCGCAGATCGAAGCGAGCTGAAGAGCGCAAGCGCGCAAGCTCAC 5961
OY 269 -----LeuGlyGluProValTyrrIleAspTyrrArgAlaValaGluValaArg1 284
Db 5960 GCTCAAGGTAGTGGCCGCGAGCGCGCTGCGCTCGCCAGGTGAATATCGAGTCTCGG 5901
OY 284 YglGluAlaAspAspLysAlaPheThrThrValaAlaAspGluValaProLeuLeuIleG1 304
Db 5900 CGAGCGCCCGCAGCTCGAGAGACTTCGCTTCGCTCGGCGCAAG---CAGCTCAAAACCGGG 5844
OY 304 yAspValaPheHisHisGlyTyrrGluThrLysAsnLeuIleGluAsnAlaSerAl 324
Db 5843 CGCGAAGCTGAACCACTGCTCTACGAGAGATGCCAGCGCTGATCCAGAACCGAGGCTTC 5784
OY 324 aGluHisGlyThrPheAspGlyArgThrPheAspArgSerAlaAspValIleLeuProAs 344
Db 5783 GCGCTACGGCTTCTCCAGAGCGCGCTTCAGCACCAAGCGCTGACATCGATCCGCGCGC 5724
OY 344 PAsnThrAlaAspValSerLeuIleTyrrAspThrGlyThrGlnTyrrArgPheAspGluVa 364
Db 5723 CGCGATCGCGCACATCGACCTGGTCTACGACAGGCGCAGCGCTACACCTTCGGCAAGGT 5664
OY 364 lValaPhePheThrIleAspProLysThrAsnGluLeuThrThrAspProAspLysLeuPr 384
Db 5663 GAGCTTC-----GACGCGGACTCGATC-- 5642
OY 384 oValLysArgLLeuLeuLeuGluGlnLeuLeuThrValaAsnMetGlyGluAlaTyrrAsnLe 404
Db 5641 -ATCGAGCAGAACTGTGGCCCGCATGCTGCTTCAGAGCGCGCCCAACCTTATGACTC 5583
OY 404 uGlnAlaValaArgAlaLeuSerAsnAspLeuIleAlaIleThrArgTyrrPheAsnMetValas 424
Db 5582 CGAAGCTATCGCGCACTCAACACAGAACTCAATCCAGCGGCTACTTC----- 5534
OY 424 nThrGluIleValaPheProGluArgGluGlnIleGlnAsnAspGlnValaSerPheGluG1 444
Db 5534 ----- 5534
OY 444 nSerSerSerSerArgThrGluProAlaGlnValaAspGluSerThrLeuGluProVal11 464
Db 5534 ----- 5534
OY 464 eGluThrValaGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAl 484
Db 5533 -----GAGGCGGTGCGGGTCGACGCCCGCCAGCCGAGCGGCGAGGC 5493
OY 484 aSerAsnLeuIleGlnAspLysLeuAsnLeuValaIalaIaLysAlaArgHisLeuTyrrAs 504
Db 5492 C----- 5492
OY 504 pMetProAspAspArgValaLeuAlaIleAsnHisAspAspGlyValaAsnArgSerIleLe 524
Db 5491 -----GACGCGCGCGCGCCAGGCGATC-- 5471
OY 524 uGlyArgIleSerAspAlaValaSerAlaValaIalaArgAlaIleLeuProAspGluSerG1 544
Db 5471 ----- 5471
OY 544 uAsnGluValaIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAs 564
Db 5471 ----- 5471
OY 564 pValTyrrGlnSerLysValaProLeuTyrrValaPheValaAlaSerAspLysProArgAs 584
Db 5470 -----CCGTAAGCGGTCCGCTGAGGCGCGCAAGCCCGGAC 5433
OY 584 pGlyGlnIleGlyLeuGlyTyrrGlySerAspThrGlyThrArgLeuValaThrLysPheG1 604
Db 5432 CATGGCGCTCGCGCTCGCTTCTCGACCGAGCTGGGGCGCGCGGAGTTCAACTGGAC 5373
OY 604 uHisAsnLeuIleAsnArgAspGlyTyrrGlnAlaGlyAlaGluLeuArgLeuSerGluAs 624
Db 5372 CCGCCATTGGGTGAATGCCGAAGACATAGCCTGGATTGCGAGGAATCTCCGACAC 5313
OY 624 pLysLysGlyValaLysLeuTyrrAlaThrLysProLeuSerHisProLeuAsnAspGlnLe 644
Db 5312 ACGGCAACACCTCGCGCTGCTGACGAGATTCCTCTGATCCGCGCTCACCGCAAGCT 5253
OY 644 uArgAlaThrLeuGlyTyrrGlnGlnGluValaPheGlyHisSerThrAsnGlyPheAspLe 664
Db 5252 GCGCTTACGACGCGCTTACCACTTCGAA-----GACCTGGTCGACAC 5211
OY 664 uSerThrArgThrLeu-----GluHisGluIleSerArgSerIleIleGlnAs 680
Db 5210 CGAAGCAGACGCTGTCACCTTGGCGGCAATGCGACAGCAAGCGCGCC----- 5162
OY 680 nGlyGlyTyrrAsnArgThrTyrrSerLysArgTyrrArgLeuAspLysLeuLysThrGlnAl 700
Db 5161 -GATGGCTGGCAGCGGGGTGCTTCTGCTGAACTGATGCGGAGAGTACAGACTG----- 5108
OY 700 aProProGluThrThrPglAspLeuProValaAspPheValaAsnGlyLysProSer----- 718
Db 5107 -----GCGCAGCATCCGGGCT 5091
OY 719 -GlnGluAlaLeuLeuAlaGlyValaIalaValaHisLysThrValaAlaAspAsnLeuValas 738
Db 5090 GAGCAGCTTCCTGATGCGGGGATCGGCTACTCGCTCGTGAGACCGCAACAAGGTCTGA 5031
OY 738 nProMetArgGlyTyrrArgGlnArgTyrrSerLeuGluValaGlySerSerGlyLeuValase 758
Db 5030 CCCCAGCATGGCTACCGCCCTGCGACGTTCAAGCTCAAGGGGCGCAAGAGGCTCTGCTGC 4971
OY 758 rAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValaTyrrSerPheGlyAspAs 778
Db 4970 CGACGCGGACTCTCCATGCGAGCCATGCGCAAGGCTTACCGACCTCGCGCGC-- 4913
OY 778 nAlaTyrrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrrIleTr 798
Db 4912 -----GGCCATCGCTGCTCGCGCGCTCGAGGTAAGGCGGAATCGC 4872
OY 798 pSerAspAsnPheAsnHisValaProTyrrArgLeuArgPhePheAlaGlyGluAspGlnSe 818
Db 4871 CACCAACGACATCAAGATCGATCCGCCCTCGCTCGCTTCTTCGCTGCGCGCGCACAGAG 4812
OY 818 rIleArgGlyTyrrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrrLeuThrG1 838
Db 4811 CGTGGCTGTACGATCAACCGAGCGCTCTTCCCGGAGATATCCGATGGCGCAGAGATCGG 4752
OY 838 yGlyGlnValaLeuAlaValaGlyThrAlaGluTyrrAsnTyrrGluPheMetLysAspLeuAr 858
Db 4751 CGGCGGTACATGATCCCGCGCAGCGTCGAGTATCAATATCCGTGGCGCGCGCTGGCG 4692
OY 858 gLeuAlaValaPheGlyAspIleGlyAsnAlaTyrrAspLysGlyPheThrAsnAspThrLys 878
Db 4691 CTTGGCAGCTTCTGTCGACGAGGCGACGCTTCAACTCGCTGAGACTTCCGCTGATCA 4632

```


[illegible]

OY	244	AlaValAlcIyTYrTyr-----AspIleAspLeuSerIlelleArgsnSerIleGlyGlu	261
Db	1590	GCAATACGGCTATTATCAAAGCGAAGTGGAATTTCAATCCAAAAGCCAAAAACTCCTAA	153
OY	262	ValAsp---ValIleIleHis----AspLeuGIglProValTyrlleaSpTYrArGla	279
Db	1530	AAGAATCACTAANTCGGCCCATGTGGAGATGGCAAACTGTGAANAATCGAGATGTGCAT	147
OY	280	ValGlValAlArgGIgLyglLylAlaAspAspLysAlaPheThrThrVaAlaAspGluval	299
Db	1470	TTTTGTCGACAGGTGAAGGCACTCAAGATCTCGAATTTTCGCCGTGACCAAGAATATC	141
OY	300	ProleuLeuIleGlyAspValPheHisHISglLYstygGluThrLysLysAsnLeuIle	319
Db	1410	CCAGAAAAA---GGAGACATCTCAATCACGAACCCTATGATAATTAATAAGTCACATC	135
OY	320	GIuAsnAlaSerAlaGlnHisglTYrPheaSpGLYArgTrPLeuAspArgSerVaAsp	339
Db	1353	CAAAAATCCCATGAGTAGTCGAGCCTATTTCGACGGTCATTTTTAGATCACCTTAGAA	129
OY	340	ValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrsphrgGlyThrGlnTy	359
Db	1293	GTCAGACCTTCAMCGACAAGCTGTGTGGCGCTTGAATATATATGTGTGAACCTTAC	123
OY	360	ArgPheaSpGluValAlPhePheThrIleaspProLysThrnsGlnLeuThrhrAsp	379
Db	1233	CGCTTCGGTGAAGTTACCTTT-----	121
OY	380	ProAspLysLeuProValLYsaRgGluLeuLeuGluGlnleuLeuthrValasnMetGly	399
Db	1212	--GAAATGCGCAAAATACGCGAAGATTATTCGTACATGTGATTAATTTAAAGG	115
OY	400	GIuAlatyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIlealThrArgTy	419
Db	1155	CAGCCCTACCTGATTATGATCTTTCTACTTAAACCAACAATTACACGTGCTCAATTGG	109
OY	420	PheasnmEvaLsnThrGIuIlleValPheProGluArgGluGlnIleGlnasnaspgLn	439
Db	1095	TTTTGGCTCTGTTTTAAATGCAACCTGTTTGTAGATGAGAACACAAATTTGTGAATGTGCAT	103
OY	440	ValSerPheglugInserSerSerArgThrGluProAlaGlnValAspGluSerThr	459
Db	1035	GTA-----	103
OY	460	LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro	479
Db	1033	-----	103
OY	480	IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAla	499
Db	1032	-----TTGCTCCAAACCAACGC-----	101
OY	500	ArgHisLeuTyraSpmetProAspAspArgValLeuAlaIleAsnHisAspAspGlyVal	519
Db	1018	-----	101
OY	520	AsnArgSerIleLeuGIyArgIIeSerAspAlaValSerAlaValAlaArgAlaIleLeu	539
Db	1018	-----	101
OY	540	ProAspGluSerGIuAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArg	559
Db	1018	-----	101
OY	560	LysThrProAlaAspValTYrgInserLysLysValProLeuTyryValPheValAlaser	579
Db	1018	-----	101
OY	580	AspLysProArgspGlyGlnIleGlyLeuGlyTYrpclYserAspPhrgIyThrArgLeu	599
Db	1017	-----AAAAAATAAGCATGAGAGGGGATTTGGCGGACATCTGATGTGGACCTCCCTTG	961

OY	267	HisAspLeu-----GlyGluProValTyrTlleAspTyrAlaGluValAlaArgGly	284
DB	5019	GCMAAGAGTGGTTCCTGGCGAGCCGGATTTGATGATCCGGCTGGATGATCTGATGGCAAGCC	5078
OY	285	Gluc1AlaAspAspLysAlaPheThrThrValAlaAspGluValProIleuIleGly	304
DB	5079	GATGCGAAGAACGGATCTCGATTATACAGGCCCTGTCGCCGGATACCCCGCAAAATTTTGGC	5138
OY	305	AspValPheHisGlyLysTyrGluThrThrLysAsnIleuIleGluAsnIleSerAla	324
DB	5139	TGGCTCTCAACCAAGGATTTTGATATTTTCACCAAGTTCTCTTACTGGTTTACCGTTA	5198
OY	325	GluHisGlyTyrPheAspGlyAlaGlyTrpLysAspArgSerValAspValIleuProAsp	344
DB	5199	CGCCGTGGCTATTTTCAGACCCCAATATGATGACAAAGCCAGCTCGGTGTCCGCTCATTA	5258
OY	345	AsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluVal	364
DB	5259	CACGACGGCTTCTGGATATTGATTTTCGACAGTGGGCAACGTTTACCGCTTGGTAAAGTT	5318
OY	365	ValPhePheThrThrIleAspProLysThrAsnGlnIleuThrThrAspProAspLysLeuPro	384
DB	5319	ATTTTC-----CAAGGCTCACAA	5336
OY	385	ValLysArgGluLeuLeuGluGlnIleuLeuThrValAsnMetGlyAlaValTyrAsnLeu	404
DB	5337	ATCCGACGAGATTTATCGCAGAAATTTAGTCCGCTTCACATAGAGGGGAAATTTATACCTCT	5396
OY	405	GlnAlaValAlaGalaLeuSerAsnAspPheIleAlaThrArgTyrPheHisMetLeuAsn	424
DB	5397	GATGACACTGGCCGAGTTGGAAACGCCGCTTGGCAGCCCAACTGGTTTACTCGGGT---	5453
OY	425	ThrGluIleValPheProGluAlaArgGluGlnIleGlnAsnAspGlnValSerPheGlnGln	444
DB	5454	-----CTTGTTCGCCGGAT-----TTCCAAAT	5477
OY	445	SerSerSerSerArgThrGluProAlaGlnValAspGlnSerThrLeuGluProAlaIle	464
DB	5478	GCCAAACGAGCTAA-----	5492
OY	465	GluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAla	484
DB	5493	-----ATTTACCGCTGAT-----	5507
OY	485	SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaIleLysAlaArgHisIleuTyrAsp	504
DB	5507	-----	5507
OY	505	MetProAspArgValLeuAlaIleAsnHisAspArgGlyValAsnArgSerIleLeu	524
DB	5507	-----	5507
OY	525	GlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGlnSerGln	544
DB	5508	-----GCGGTGCTTACCCACGACGCTGAG	5531
OY	545	AsnGluValIleAspLeuProGluAlaGlyThrAlaLeuAlaAsnArgLysThrProAlaAsp	564
DB	5532	AATACCGCTC-----	5540
OY	565	ValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAsp	584
DB	5540	-----	5540
OY	585	GlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuValThrLysPheGln	604
DB	5541	---GACTCGCGGCGGCTTATCGCAACCGATGTCGCTCCGCTCTGACCGCAGTGGCGCT	5597
OY	605	HisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGlnAsp	624
DB	5598	AAGCTTGGATGAAGTCTCTTGGCGCATACCTTTAACCAACCACTAACGCGCTTGGCGACT	5657

[illegible]

RESULT 8	
AJ414157/c	
LOCUS	216050 bp DNA linear BCT 06-JUN-2002
DEFINITION	Yersinia pestis strain C092 complete genome; segment 17/20.
ACCESSION	AJ414157 AL590842
VERSION	AJ414157.1 GI:15981328
KEYWORDS	
SOURCE	Yersinia pestis.
ORGANISM	Yersinia pestis
	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
	Yersinia.
REFERENCE	1 (bases 1 to 216050)
AUTHORS	Parkhill,J., Wren,B.W., Thomson,N.R., Titball,R.W., Holden,M.T.G.,

	/note="Pfam match to entry PF02261 Asp_decarbox, Aspartate decarboxylase, score 264.30, E-value 1.6e-75"
gene	complement(1880. .3073) /gene="YP03404" complement(1880. .3073) /note="Similar to Escherichia coli hypothetical protein Yade SW:YADE_ECOLI (F91666) (409 aa) fasta scores: E(): 0, 65.4% id in 382 aa"
CDS	/codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="CAC92634.1" /db_xref="GI:15981331" /db_xref="SPTREMBL:Q82BK5" /translation="MTTERDSEVSLGVEHVIIPGEIKREGOLIQTVPFLVIAEYEFKFG NCTGISDDDLREINKAQNDMDGLDKRPINONILMRPAKYLAADRSDEDFAL AENLEPYIGKLKRLSTNTWYNVINIGDELGYVSADAIEDNGIPILYHHHLKEENR RLNSTSTTSQVAFSNOMAYLKAGQVDYISYLEEGYLNNKINLPAAVVVLFDGGLK SVSRVAPYIKENGFRATFIISRIRKHPPKMDSLOEPSIELKGIDVDPIOSH THPLRTNRHRHPILLSSRNHIIFDRPSRRALOAFNPAPHILSYSPGNCRANVNA AKNAEFLHAVTTWGCKVAPGDNPITLKRLLTIRDISTPMERIANOPGVAAKAPADA VLEM"
gene	complement(3400. .4170) /gene="YP03405" /note="yadh" complement(3400. .4170) /gene="YP03405" /note="Similar to Escherichia coli hypothetical 28.5 kDa protein in hpt-pand intergenic region Yadh SW:YADH_ECOLI (P36880) (256 aa) fasta scores: E(): 0, 80+1% id in 256 aa, gkud2 to Vibrio cholerae putative permease V6059 TR:GOKUD2 (EMBL:AE004144) (256 aa) fasta scores: E(): 0, 66.0% id in 256 aa"
CDS	/codon_start=1 /transl_table=11 /product="putative ABC transporter, integral membrane protein" /protein_id="CAC92635.1" /db_xref="GI:15981332" /db_xref="SPTREMBL:Q82BK4" /translation="MTRLIYHALOSIWKETTRRAKIOTLVPPVTMSLYFTSEN LTGAIGMGGRDIQMOTIVPGLLIMAVITNAYSVAASFCAKFORSIEELVAAPVT HLVILGYGGVGARGICGILVTIIISLFEPVLSHVSMMLTLITALLIFSGLGLN AFEATFPDISLAPTEFLVPLTYLGVFYSLSLPPFEWAASKLNPIVMISGRYGF LGITDVSLATRTGVLVPIAVEYAWAWLIERGGLRT"
misc_feature	complement(3409. .4164) /gene="YP03405" /note="Pfam match to entry PF01061 ABC2_membrane, ABC-2 type transporter, score -22.30, E-value 5.8e-05"
misc_feature	complement(3427. .3492) /gene="YP03405" /note="one of 6 probable transmembrane helices predicted for YPO3405 by TMHMM2.0"
misc_feature	complement(3526. .3639) /gene="YP03405" /note="one of 6 probable transmembrane helices predicted for YPO3405 by TMHMM2.0"
misc_feature	complement(3579. .3657) /gene="YP03405" /note="one of 6 probable transmembrane helices predicted for YPO3405 by TMHMM2.0"
misc_feature	complement(3679. .3744) /gene="YP03405" /note="one of 6 probable transmembrane helices predicted for YPO3405 by TMHMM2.0"
misc_feature	complement(3790. .3855) /gene="YP03405" /note="one of 6 probable transmembrane helices predicted for YPO3405 by TMHMM2.0"
misc_feature	complement(3919. .3984) /gene="YP03405" /note="one of 6 probable transmembrane helices predicted for YPO3405 by TMHMM2.0"
misc_feature	complement(3919. .3984) /gene="YP03405"

	misc_feature	/for yP03405 by TMHMM2.0" complement(4045..4110) /gene-"yP03405"
	gene	/note-"One of 6 probable transmembrane helices predicted for yP03405 by TMHMM2.0" complement(4167..5093) /gene-"yP03406" /note-"yadG" complement(4167..5093) /gene-"yP03406"
CDS		/note-"Similar to Escherichia coli hypothetical ABC transporter ATP-binding protein YacD SW:VADG_ECOLI1 (P36873) (308 aa) fasta scores: E(): 0, 81.8% id in 308 aa, and to Vibrio cholerae ABC transporter ATP-binding protein VC0589 TR:O9KUDJ (EMBL:AEO04144) (305 aa) fasta scores: E(): 0, 71.4% id in 304 aa" /codon_start=1 /transl_table=1 /product="putative ABC transporter, ATP-binding protein" /protein_id="CAC92636.1" /db_xref="GI:1598133" /dx_xref="SPTREMBL:O8ZBK3"
	misc_feature	/translation="MTYALEITDTLTNYMGVOALRGIDLVHADGFVALCPNGNA STTGISLSLVNKTSGKVDFEVDINDNDVNARQGLVPSEFNEPFETVLQIVY AGPYGVTRDALRAEKYLSQLDMWRDERAIRISGMKRRLMARLMHEPKLL DEPVAGVIDLARSMVGFLKEINAOCTTIILTHYLEAEMLCRNIGLIIONELVE TMKOILLKETSETFFFDLGAKSPLEKGGRYLTTSLVEDVKRQGCLNSLFSSO VCGOVOSMRKANRKARELPDLTVLNHGEBKA"
	misc_feature	complement(4458..5003) /gene-"yP03406"
	misc_feature	/note-"Plam match to entry PFO0005 ABC_tran, ABC transporter, score 198.80, E-value 8.5e-56" complement(4641..4685)
	misc_feature	/gene-"yP03406"
	misc_feature	/note-"PS00211 ABC transporters family signature." complement(4959..4982) /gene-"yP03406"
gene		/note-"PS00017 ATP/GTP-binding site motif A (P-loop)." 5329..5591 /gene-"yP03407"
		/note-"yadF" 5329..5591 /gene-"yP03407"
CDS		/gene-"yP03407"

Alignment Scores:	Pred.	No.:	Length:	Matches:	Mismatches:	Gaps:
Scores:	Score:	494.00	216050	167	105	17
Percent Similarity:		38.04%		Conservative:	255	
Best local Similarity:		23.36%		Mismatches:	188	
Query Match:		10.45%		Indels:	17	
IDB:		1				

US-09-914-168-2 (-1-919)	x AJ414157 (-1-216050)
Oy	211 Asnileysalaalalegluaspile-----ThcInguIsusefalametAspleuaSngly 229
Dd	136908 AACGCACGGCAGCAgcttttccAcAATTgcAgcAtTAgaATTAACCggAGAT-----GCC 136855
Oy	230 Serle---ProatrgleunargclnthralaleuValAlaalArgalavaIgltYrTYr 248
Dd	136854 cgttttcctTCACgggtcgATAagAACgATTAATCCGcAGGCGTGCGCcCATTAGGTATTAC 136795
Oy	249 Asp-----lleasplueserillelaeargasnerillegluvalasvallielle 266
Dd	136794 CATCCCACCAAttaccgattGAATTCAAAACCGCCCCGCCAccCGCActTtagtgTTAAATT 136735
Oy	267 Hlasplen-----glyluaprovaItyrIlleaSpTyrafgalavalaglualaTarqly 284
Dd	136734 GCAAAGAAGTgctTCCtgGGAGACCcGGrATTgATTTgCCGtGCATATCGACTGCAAGGC 136675
Oy	285 GlugylalasapaspLyalsaphethrThValAlaslaspeluValproleuleulliegly 304
Dd	136674 gGTGCGAAAAACGgATTCCTGATTAATTCACCAcCATTCGTCGGCGGTATGATACCCGAAAAATTGGC 136615

OY	305	AspValPheHisIleGlyLysTyrGlnThrLysLysAsnIleGluAsnAlaSerAla	324
Db	136614	TCGGTCTCAACCATGTGATTTGATTAATTCACCAAGTTCTCTACTGGTTAAGCGTTA	1365555
OY	325	GluHisGlyTyrPheAspGlyArgTrpLeuAspArgSerValAspAlaIleLeuProAsp	344
Db	136554	CGCGGTGCTATTTCGACGCCAATATGATCAAAAAGCCAGCGTGGGTATGGCGGTCAATTA	1364955
OY	345	AsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrGlnPheAspGluVal	364
Db	136494	CACGAGCGGTTCTGGGATATTGATTGATTCGACAGTCGGGCAACGTTACCCCTTGGTAAAGGTT	1364355
OY	365	ValPhePheThrIleAspProLysThrAsnGlnLeuThrAspProAspLysLeuPro	384
Db	136434	ATTTC-----CAAGCTCAACA	136417
OY	385	ValLysArgGluLeuGlnGlnLeuLeuThrValAsnMetGlyAlaTyrAsnLeu	404
Db	136416	ATCCGAGAGGATTTATCTGCACAAATTTAGTCGCTTCATAGAGGGGAATCTATTACCTCT	1363575
OY	405	GlnAlaValArgAlaIleSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsn	424
Db	136356	GATGACATGGCCGAGATTGAACCCCGCTGGCACCACTGGTTAATCTGGGTG---	1363000
OY	425	ThrGlnIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGln	444
Db	136299	---GTGTTTCGCGGAT---TTCCAAAGAT	136276
OY	445	SerSerSerSerArgThrGlnProAlaGlnValAspLysThrLeuGlnProValIle	464
Db	136275	GCCCAAGAGATTAA-----	136261
OY	465	GluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAla	484
Db	136260	-----ATTTCACCGTTGAT-----	136246
OY	485	SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp	504
Db	136246	-----	136246
OY	505	MetProAspArgValIleuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeu	524
Db	136246	-----	136246
OY	525	GlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGlu	544
Db	136245	-----CGGTCTGTTACCCACGACCTAGAG	1362222
OY	545	AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAsp	564
Db	136221	AATACCGCT-----	136213
OY	565	ValTyrGlnSerLysValProLeuTyrValPheValAlaSerAspLysProArgAsp	584
Db	136213	-----	136213
OY	585	GlyGlnIleGlyLeuGlyTyrPglySerAspThrGlyThrArgLeuValThrLysPheGlu	604
Db	136212	---CAACGGGGGGCGTTATGCAACGATGCGTCCGCTGACCGCCACCTGGCGCT	136156
OY	605	HisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlnLeuArgLeuSerGlyuAsp	624
Db	136155	AAGCCTTGATGAACCTCCTTCGGGATATGCTTAACCAACCACTACCGCGCTTTCGGCACCT	136096
OY	625	LysLysGlyValLysLeu---TyrAlaThrLysProLeuSerHisProLeuAsnAspGln	643
Db	136095	GAGCAGACACTGATTTTCAGCTACCGTATTCCTTATTGAGAAACCCCTCGAGCAGATAT	136036
OY	644	LeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAsp	663
Db	136035	TACCTGATCAAGGGCGCTTTAAGCGT-----ACCGACTTAACGAC	135994

```

LINE 21534948
UMED 11677609
RENCE 2 (bases 1 to 24892)
THORS The Salmonella typhimurium Genome Sequencing Project.
TE Direct Submission
URL Submitted (29-MAR-2001) Genome Sequencing Center, Department of
Genetics, Washington University School of Medicine, 4444 Forest
Park Boulevard, St. Louis, MO 63108, USA
COMMENT Supported by NIH grant 50 01 A143283

Coding sequences below are predicted from manually evaluated
computer analysis, using similarity information and the programs;
GLIMMER: http://www.tigr.org/softlab/glimmer/glimmer.html and
Genemark: http://opal.biology.gatech.edu/Genemark/
EC numbers were kindly provided by Junko Yabuzaki and the Kyoto
Encyclopedia of Genes and Genomes: http://www.genome.ad.jp/kegg/,
and Pedro Romero and Peter Karp at EcoCyc:
http://ecocyc.org/ecocyc.com/ecocyc/

The analyses of ribosome binding sites and promoter binding sites
were kindly provided by Heladia Salgado, Julio Collado-Vides and
Reguiondb:
http://kinich.cifn.unam.mx:8850/db/reguiondb\_intro.frameset

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistries or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one m3 subclone.

Location/Qualifiers
1. 24892
/organism="Salmonella typhimurium LT2"
/strain="LT2: SGSC 1412; ATCC 700720"
/db_xref="ATCC:700720"
/db_xref="taxon:99287"
/note="LT2"
complement(95. .383)
/gene="yjfY"
/note="STM4389"
complement(95. .370)
/gene="yjfY"
/note="yjfY"
/note="similar to E. coli orf, hypothetical protein
(AAC77156.1): Blastp hit to AAC77156.1 (91 aa), 718
identity in aa 1 - 91"
/codon_start=1
/transl_table=11
/product="putative outer membrane protein"
/protein_id="AAL23209.1"
/db_xref="GI:16422951"
/transl_table="MKRIMLEFLAALLPGITATAVAELNNHQRNMDVRSLGVIYIN
HNFATESEANLALNEADVVRNMYHVLIREPGSNGNIHSAANIYR"
complement(378. .383)
/gene="yjfY"
/note="putative RBS for yjfY; RegulonDB: STMSIH004274"
560. .770
/gene="STM4390"
560. .565
/gene="STM4390"
/note="STM4390"
/note="putative RBS for STM4390; RegulonDB: STMSIH004275"
564. .770
/gene="STM4390"
/note="hypothetical protein"
/codon_start=1
/transl_table=11
/product="putative cytoplasmic protein"
/protein_id="AAL23210.1"
/db_xref="GI:16422952"
/transl_table="MKOKSALTVISLPGQAVYSGMREYVNISSAKCDGTGTFGTTI
NCTTTLTSLPAGASITTRHFSADL"
828. .1236
/gene="rfpF"

```

```

RBS      /note="STM4391"
          /gene="rpsf"
          /note="putative RBS for rpsf; RegulonDB:STWSIH004276"
          828..833
CDS      /gene="rpsf"
          /note="similar to E. coli 30S ribosomal subunit protein S6
          (AAC77157.1); Blastp hit to AAC77157.1 (131 aa), 99%
          identity in aa 1 - 131"
          /codon_start=1
          /transl_table=11
          /product="30S ribosomal subunit protein S6"
          /protein_id="AL23211.1"
          /db_xref="GI:16422953"
          /translation="MRHVEIVFMVHPDSEQVGMIERYSAAITGAEGKIHRLDWMGR
          VOLAYPIKRLKHAHYLVNVEAPQEVIDEITTFRENDVAISMVRTKHAVTEASPM
          VKADRRERRRDFANETADDAAGDSEE"
          1226..1557
gene      /gene="prlB"
          /note="STM4392"
          1226..1231
RBS      /gene="prlB"
          /note="putative RBS for prlB; RegulonDB:STWSIH004277"
          1243..1557
CDS      /gene="prlB"
          /note="similar to E. coli primosomal replication protein N
          (AAC77158.1); Blastp hit to AAC77158.1 (104 aa), 99%
          identity in aa 1 - 104"
          /codon_start=1
          /transl_table=11
          /product="primosomal replication protein N"
          /protein_id="AL23212.1"
          /db_xref="GI:16422954"
          /translation="MTNRLALSGTVCAPLRKVPSGIPHCQVLEHRSVQEEAGFHR
          QAMQMPVIVSGHENQAITHSITVGSRTIVOGFISCHKAKNGLSKMWLHAEOIELIDS
          GD"
          1547..1789
gene      /gene="rpsr"
          /note="STM4393"
          1547..1552
RBS      /gene="rpsr"
          /note="putative RBS for rpsr; RegulonDB:STWSIH004278"
          1562..1789
CDS      /gene="rpsr"
          /note="similar to E. coli 30S ribosomal subunit protein
          S18 (AAC77159.1); Blastp hit to AAC77159.1 (75 aa), 98%
          identity in aa 1 - 75"
          /codon_start=1
          /transl_table=11
          /product="30S ribosomal subunit protein S18"
          /protein_id="AL23213.1"
          /db_xref="GI:16422955"
          /translation="MARVPRRRFCRFTAEQVQELDKDIATLKNYTESGKIVPSRI
          TGTAKVOROLRAIKRAHYLSILPYTDHQQ"
          1819..2280
gene      /gene="rplI"
          /note="STM4394"
          1819..1824
RBS      /gene="rplI"
          /note="putative RBS for rplI; RegulonDB:STWSIH004279"
          1831..2280
CDS      /gene="rplI"
          /note="similar to E. coli 50S ribosomal subunit protein L9
          (AAC77160.1); Blastp hit to AAC77160.1 (149 aa), 95%
          identity in aa 1 - 149"
          /codon_start=1
          /transl_table=11
          /product="50S ribosomal subunit protein L9"
          /protein_id="AL23214.1"
          /db_xref="GI:16422956"
          /translation="MQVILIDKVANLGSIGDOVNVAGYARNEFLVPEKGAVPATKKNV
          EYFEARRAELAEKLVADLAANARAERKIALVELVTIASAGDEGKLFJSGTRADIADA
          VTAAGVDVAKSEVRLPNCVLRITGGEHVFQVHSEVFATIIINVAAE"
```

```

gene      2417..3354
          /gene="yifZ"
          /note="STM4395"
          2417..2422
RBS      /gene="yifZ"
          /note="putative RBS for yifZ; RegulonDB:STWSIH004280"
          2428..3354
CDS      /gene="yifZ"
          /note="similar to E. coli orf, hypothetical protein
          (AAC74606.1); Blastp hit to AAC74606.1 (266 aa), 25%
          identity in aa 20 - 250"
          /codon_start=1
          /transl_table=11
          /product="putative permease"
          /protein_id="AL23215.1"
          /db_xref="GI:16422957"
          /translation="MDTORQASPFARKNVVYCAACCLIMGSAVPAIKSGYDLPOIA
          TDDIPSKIYFACGRFLPAGGLLFLPLDRKPIGRPRRPFQRODTLLGLTQSLQTLF
          FYIGLAFITSGVKGKIMNATGTFEVSLLAHFTYQNDRLSYNKTGLGIFAGVNVNVS
          NGLDFSPNLPBGSSVLAALISAATLVGRKLSQTDVPMVTGYQIGGLVLYIGY
          VFGGLTILHGFSSVAITLVLTLLSVAFALMSITLKYNRVGMIAPENFLIPVGAALS
          AIFIGENILEMKYMLVLCSGIMVWNVKR"
          complement(3404..4052)
          /gene="yrfB"
          /note="STM4396"
          complement(3404..4039)
CDS      /gene="yrfB"
          /note="similar to E. coli orf, hypothetical protein
          (AAC77163.1); Blastp hit to AAC77163.1 (224 aa), 85%
          identity in aa 13 - 224"
          /codon_start=1
          /transl_table=11
          /product="putative cell envelope opacity-associated
          protein A"
          /protein_id="AL23216.1"

Alignment Scores:
Score: No.: 2,28e-20 Length: 24892
Percent Similarity: 486.50 Matches: 174
Best Local Similarity: 36.90% Conservative: 112
Query Match: 22.45% Mismatches: 270
DB: 10.29% Indels: 219 Gaps: 19

US-09-914-168-2 (1-919) x AE008906 (1-24892)
Qy 172 Leupheasnaspcgiyalaaslnys-----ValProArgLeuLysAla 185
      ::::: ||| |||:::
Db 16338 ATTTTCCAGGGGTGGGATATGACGAGAAATGCGCACATATCCGTCAG 16397
Qy 186 LysPheTyrglnSerSerglnSerglyGluThrSerAlaIleGlySerHisGlnLys 205
      ::::: ||| :::::
Db 16398 TTATGTTGGGTTAGCTACGTACGTGCTGACGACGATCTCGGGTAGCC----- 16442
Qy 206 ThrGluProTyrlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGlnSerAlaMet 225
      ::::: ||| |||:::
Db 16443 -----GCAAATGTTGCTGTAAGAGCGATATCCGAGAGCTGCAAAA 16490
Qy 226 AspleuasnglyserIleProArgLeuArgGlnThrAlaLeuValaAlaArg----- 243
      ::::: ||| |||:::
Db 16491 AACGTTGCGGCACAGCTCTAGCATTCACAGCAGATGAGGTGACCGCGGCGCTTT 16550
Qy 244 -----AlaValGlyTyTyraSP--- 249
      ::::: ||| |||:::
Db 16551 CGGCCCGAGCTGATGATTCGATTCGCGAAGGCGTTAAACGCTTAGCTATTCGAACCC 16610
Qy 250 ---IleAspleuserIleIleArgAsnSerIleGlyGluValAspValIleIleHisAsp 268
      ||| :::::
Db 16611 ACTATCAAAATTGCATCTGCTTCACCGCGCTGCGAAAGACGCGAGTATTAATCGCCAGG 16670
Qy 269 Leu-----GlyGluProValTyrlleAspTyraArgAlaValGluValArgGlyGluGly 286
      ::::: ||| |||:::
Db 16671 GTTACGCCCGGCAGCGGCTGATTGGCGGCTACAGAGATCATTTTACCGCGGCGCGG 16730
```


OY	287	AlAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal	306
Db	16731	CGCAGCGAAAGAAATTAATGCGCTTACTGAAAGCGCTCG--GCAATTGGCAGGGG	16787
OY	307	PheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerIleGluHis	326
Db	16788	CTGACACGGGAGCATATGACAAATTTTAAAGTGGTTAACGAGGTGGCTGACGTAA	16847
OY	327	GlyTyrPheAspGlyArgTrpLeuAspArgSerValAspValIleLeuProAspAsnThr	346
Db	16848	GGCTATTTCGACGCGAATCTTAAAGTCAAGTTGGCAATTCCTGTGGCGCATCG	16907
OY	347	AlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhe	366
Db	16908	CGGTTTGGGATTAATGATTATGATAGCGGTAGCGCTATCGCTTGGACCTTCACCTTC	16967
OY	367	PheThrIleAspProLysThrAsnGluLeuThrThrAspProAspLysLeuProValLys	386
Db	16968	-----GAAGGTTTCGACATTCGT	16985
OY	387	ArgGluLeuGluGluLeuLeuThrValAlaAspGluValGluAlaTyrAsnLeuGlnAla	406
Db	16986	GATGAATATTTACAAATCTGTGGCCGTTTAAAGAGGGTGATGACATCGAAAGAT	17045
OY	407	ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGlu	426
Db	17046	TTGGCGGAACCTAAACCGCGCATTTCCGCTACCGGATGTTAATTCGTGTC-----	17096
OY	427	IleValIlePheProGluArgGluGlnIleGlnAsnAspGluValSerPheGluLeuIleSer	446
Db	17097	GGTGGGCGCCGGA-----TTTGAGAAATCGCGT	17126
OY	447	SerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThr	466
Db	17127	AAAAAATAATCTCGCG-----	17144
OY	467	ValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsn	486
Db	17144	-----	17144
OY	487	LeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetPro	506
Db	17144	-----	17144
OY	507	AspAspArgValLeuAlaIleAsnHisAspAspGlyValAlaAspArgSerIleLeuGlyArg	526
Db	17144	-----	17144
OY	527	IleSerAspAlaValSerIleValAlaArgAlaIleLeuProAspGluSerLysGlnGlu	546
Db	17145	-----TTAAAGCGGTGTTCGCCCGGGAACCGAAGAACAG	17180
OY	547	ValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr	566
Db	17181	ATT-----	17183
OY	567	GlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGln	586
Db	17184	-----GAA-----	17186
OY	587	IleGlyLeuGlyTyrProLysAspThrGlyThrArgLeuValThrLysPheGluHisAsn	606
Db	17187	ACCGGGGTGGGTACTCCACCGACGTCGGGGCGCGGTAAAGCGTCGTGGAAAAACCG	17246
OY	607	LeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLys	626
Db	17247	TGGATGAACTCTTACGGTCATAGCCGACACACAGCAGACAGATTTCTCGCGGAAACAG	17308
OY	627	GlyValLysLeu---TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArg	645
Db	17307	GTGCTGGATTTTAGTTATAAAATGCCGCTGTAAGAAAAACCGCTGGAGCAATACTATCTG	17366

QY	646	AlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyAspLeuSer	665		
QY	17367	GTACAGGGCGGTTTAAACGTAC-----GATTGAAGATACGAGGAGAC	17414		
QY	666	ThrArgThrLeuGlnHisGlyIleSerArgSerIleIleGlnAsnGlyTyrTrpAsnArg	685		
Db	17415	TGCACCCAGCGTT-----GCCGTGCGGGTTACGTGGACCTTTCCAGTGGATTGCCAGCGC	17468		
QY	686	ThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrp	705		
Db	17469	CGGATTAACTTCGCGTGGAGTTTTCGAC-----	17495		
QY	706	GlnAspLeuProValAspPheValAsnGlyProSerGlnGluAlaLeuLeu-----	723		
Db	17496	-----CACTTTTACCGAGGGGAACGTCCACCATACCAACCATGCTGTTTTCAC	17540		
QY	724	AlaGlyValAlaValAlaHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyr	743		
Db	17541	CCGGCGCGTGAATATACGGCGAACCCTGCGCGAGGGGAGTACATGCCAACCTGGGGCGAT	17600		
QY	744	ArgGlnArgTyrSerLeuGlnValAlcLysSerSerGlyLeuValSerAspAlaAsnMetAla	763		
Db	17601	TCCGACGCGTATTCAGTAGACATATTCGATACAGCCCTCGGGCTCCGACGTGATTTCTCC	17660		
QY	764	IleAlaArgAlaGlyIleSerGlyValLysSerPheGlyAspAsnAlaTyrGlySerAsn	783		
Db	17661	GTCGTCAGGCGCCAAACGTCGTGGATTCGAGATTTGTACGATCCG-----	17705		
QY	784	ArgAlaHisGlnMetThrGlyLysIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsn	803		
Db	17706	-----CACCGTTTGTGATCGCGGCCCAATTGGGCTGGATATGAACCGCGATTTTCAC	17759		
QY	804	HisValProTyrArgLeuArgPhePheAlaGlyLysAspGlnSerIleArgGlyTyrAla	823		
Db	17760	AAAGTTCGCGCCGATTTACGTTCTTCCCGCGGGCGACCCGACATTCGCGGCTATTA	17819		
QY	824	HisAspSerLeuSerProIleSerAspLysGlyTyrTrpLeuThrGlyGlnValLeuAla	843		
Db	17820	TTCAAATCATTTCCCTTAAGATAGCGACGCGCAATTTCAAAGCGCCCTCAAAACTGGCA	17879		
QY	844	ValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGly	863		
Db	17880	ACCGATGCGCTGAGACTACCATATACGTACACCGGTAATGTTGGGGGCGAGTGTTC	17939		
QY	864	AspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyAl	883		
Db	17940	GATACGCGCGAGCGCGTAGATATTCGCGCTGACGATTCACAAACCGGCGACGGGTC	17999		
QY	884	GlyValArgTyrPheAlaSerProValGlyGlnValAlaGlyValAsp-----ValAlaThrGly	901		
Db	18000	GGCGTCGCGTGGCGCTCCCGGTTGGGCTGCAACATTCGATTTGCCGTACCGGTCGCG	18059		
QY	902	ValLysGlnGluGlyAsnProIleLysLeuHisPhePheIleGly	916		
Db	18060	GACAAAGACGACACCGGT-----TTACAGTTTATATCGGT	18095		
RESULT 10	AL627283	249050 bp	DNA	linear	BCT 06-JUN-2002
LOCUS	AL627283				
DEFINITION	Salmonella enterica serovar Typhi (salmonella typhi) strain CT18,				
ACCESSION	AL627283	AL513382			
VERSION	AL627283.1	GI:16505370			
KEYWORDS					
SOURCE	Salmonella enterica subsp. enterica serovar Typhi.				
ORGANISM	Salmonella enterica subsp. enterica serovar Typhi				
REFERENCE	1 (bases 1 to 249050)				
AUTHORS	Parkhill,J., Dougan,G., James,K.D., Thomson,N.R., Pickard,D.,				
	Wain,J., Churcher,C., Mungall,K.L., Bentley,S.D., Holden,M.T.G.,				
	Sebaiha,M., Baker,S., Basham,D., Brooks,K., Chillingworth,T.,				
	Comerton,P., Cronin,A., Davis,P., Davies,R.M., Dowd,L., White,N.,				

[illegible]

QY	627	GLYALysLeu-----lylAlThrLysProLeuSerHisProLeuAsnAspGlnLeuArg	645
Db	154086	GTCTGATTTAGTTAGTAAATGCGGCTGTAAGAACCCGCTGGACCATTAATCTG	154145
QY	646	AlaThrLeuGlyGlnGlnGlnValPheGlyHisSerThrAsnGlyPheAspLeuSer	665
Db	154146	GTACAGCGGGTTTAAAGTAC-----GATTGCAATGATACGACACAGAC	154193
QY	666	ThrArgThrLeuGlnHisGlnLysSerArgSerIleIleGlnAsnGlyLysPheAsn	685
Db	154194	TCGACACGCTT-----GCCCTGCGGTTACTGCGACCTTCCACGCGTTGGACGCG	154247
QY	686	ThrTyrSerLeuArgTyrArgLeuAspLysLeuThrGlnAlaProProGluThrTyr	705
Db	154248	GCGATTAACTGCGCTGGAGTTTCAC-----	154274
QY	706	GlnAspLeuProValAspPheValAsnGlyLysProSerGlnGlnAlaLeu-----	723
Db	154275	-----CACTTACCCAGGCAACCTCACCATACCAACGCTGTTCTAC	154319
QY	724	AlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyr	743
Db	154320	CCGGCGTGAATGATCAGCCGCAACCCGCTCGGAGGCGGATGATGCCAACCTGGGCGAT	154379
QY	744	ArgGlnArgTyrSerLeuGlnValGlySerSerGlyLeuValSerAspAlaAsnMetAla	763
Db	154380	TCGCAGCCCTACTCGTAGACTATTGCATACGGCCCGGGCTCCGACCTGATTTCTTC	154439
QY	764	IleAlaArgAlaGlyLysSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsn	783
Db	154440	CTACTGCGAGCGGCAAAACGCTCGGATTCGACTTTGTACGATTCG-----	154484
QY	784	ArgAlaHisLysLeuThrGlyGlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsn	803
Db	154485	-----CACCGTTTGTGATGCGCGCAATTGGCTGATGAAACGCGTATTTCGAC	154538
QY	804	HisValProTyrArgLeuArgPhePheAlaGlyLysAspGlnSerIleArgGlyTyrAla	823
Db	154539	AAAGTTCGCGCGGATTTAGCTTTCTTCCGCGGCGGACCGACGATTCGCGGCGATAA	154598
QY	824	HisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAla	843
Db	154599	TACAAATCTATTTCGCTAAAGATGACGCGCAATCTTAAAGGCGCTCAAAACCTGCA	154658
QY	844	ValGlyThrAlaGlnLysAntyrGluPheMetLysAspLeuArgLeuAlaValPheGly	863
Db	154659	ACCGCATTCCTGAGTACAGTAAACGTACACCGCAAAATGCTGGGCGCGCTTTTCTC	154718
QY	864	AspIleGlyAsnAlaTyrLysPheGlyPheHisAspThrLysIleGlyAlaGlyVal	883
Db	154719	GATAGCGCGGACGGCTTAGTATTTGCGCGTACGCGATTTCAAAACCGGACCGGCTC	154778
QY	884	GlyValArgTyrPheLysProValGlyValAlaArgValAsp-----ValAlaThrGly	901
Db	154779	GCGCGCGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG	154838
QY	902	ValLysGlnGlnGlyAsnProIleLysLeuHisPheIleGly	916
Db	154839	GACAAAGACAAACACGCT-----TTACAGCTTTATATACGT	154874
RESULT 11			
LOCUS	AE004323/c	14390 bp	DNA linear BCT 31-JUL-2000
DEFINITION			Vibrio cholerae chromosome I, section 231 of 251 of the complete
ACCESSION	AE004323		chromosome.
VERSION	AE004323.1		
KEYWORDS	GI:9657119		
SOURCE			Vibrio cholerae.
ORGANISM			Vibrio cholerae
REFERENCE			Bacteria: Proteobacteria: gamma subdivision: Vibrionaceae: Vibrio.
AUTHORS			Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L.,

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS
		Nature 406 (6795), 477-483 (2000)	20406833	10952301	2 (bases 1 to 14390)
					Heidelberg, J.F., Eisen, J.A., Nelson, W.C., Clayton, R.A., Gwinn, M.L.,
					Dodson, R.J., Haft, D.H., Hickey, E.K., Peterson, J.D., Umayam, L.,
					Gill, S.R., Nelson, K.E., Read, T.D., Tettelin, H., Richardson, D.,
					Ermoiaeva, M.D., Vamathevan, J., Bass, S., Qin, H., Dragol, I.,
					Sellers, P., McDonald, L., Uitterback, T., Fleishmann, R.D.,
					Nierman, W.C. and White, O.,
					DNA sequence of both chromosomes of the cholera pathogen Vibrio
					cholerae
TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS
					Submitted (14-JUN-2000) The Institute for Genomic Research, 9712
					Medical Center Dr. Rockville, MD 20850, USA
					Location/Qualifiers
					1. .14390
					/organism="Vibrio cholerae"
					/strain="N16961"
					/serotype="O1"
					/db_xref="taxon:666"
					/chromosome="1"
					/note="biotype: El Tor"
					80. .661
					/gene="VC2536"
					80. .661
					/gene="VC2536"
					/note="identified by glimmer2; putative"
					/codon_start=1
					/transl_table=1
					/product="conserved hypothetical protein"
					/protein_id="AAP5677.1"
					/db_xref="GI:9657120"
					/translation="MKNNARKNQKAPWEEIEIIVSRTLEKNDMLALOKIGELVW
					LKPSALKFLPEDLAEIKDQAEKNEARROLOQIKIMHIDPEFLALDLKLR
					KHSQTALHKLRLDRLVAEDSAIEVAMQYEDARORRLRLARASKEKKNKPR
					PKSREIFOLKELAMLAKEIEEESDDLSAE"
					complement (764. .1498)
					/gene="VC2537"
					complement (764. .1498)
					/gene="VC2537"
					/note="similar to GB:D10483 SP:P31548 PID:216486 GB:U00096
					PID:I786253; identified by sequence similarity; putative"

gene	6334 . .6597	/gene="VC2543"	6334 . .6597	/note="identified by Glimmer2; putative"	/codon_start=1	/transl_table=1	/product="hypothetical protein"	/protein_id="AA95684.1"	/db_xref="GI:9657127"	/translation="MNVHKNPALFLCDWNGSIIPIRYAKASAAVQGVNSKMSITIKVRSPSVSITPVKRSKRYNSLQTRCIAQSIPSEDPITLLPFPALH"	651. .7661	/gene="VC2544"	651. .7661	/gene="VC2544"	/note="similar to SP:P09200 GB:X12545 PID:41416 PID:537074 GB:U00096; identified by sequence similarity; putative"	/codon_start=1	/transl_table=1	/product="fructose-1,6-bisphosphatase"	/protein_id="AA95685.1"	/db_xref="GI:9657128"	/translation="MOKMRTLGEFIVKOHDFPHASGELSILASTIRLAKIVYREINIKAGIDITAGSNDNIOGEEFOOKDITVANKERFALALRQCVASAEDEAVAFSKELNNKAYVYLMPLDSSNIDVAVSGTIFSTIRRPSPGTPPTQEDFIQPGKQVLAAGVITGSSTMLVTTGNGVNGFTYDLSGFTFYSHNMRIPEMGKLYTSINEGATIRFPTGKRYKIRFCQENVPEEGRPYTSRYIGSLVADHRNLKGGITLYLPSYTSHPNGKLRLLYCNPMAPFLIEQAGGLASDGARRIMDIKPTLELHQRPPEFVSGKNMVHKEFLETYPD"	7831 . .8361	/gene="VC2545"	7831 . .8361	/gene="VC2545"	/note="similar to GB:U14003 SP:P17288 PID:450373 PID:537068 GB:U00096; identified by sequence similarity; putative"	/codon_start=1	/transl_table=1	/product="inorganic pyrophosphatase"	1.24e-18	Length: 14390	Matches: 164	Conservative: 104	Mismatches: 238	Indels: 226	Gaps: 20
gene	6334 . .6597	/gene="VC2543"	6334 . .6597	/note="identified by Glimmer2; putative"	/codon_start=1	/transl_table=1	/product="hypothetical protein"	/protein_id="AA95684.1"	/db_xref="GI:9657127"	/translation="MNVHKNPALFLCDWNGSIIPIRYAKASAAVQGVNSKMSITIKVRSPSVSITPVKRSKRYNSLQTRCIAQSIPSEDPITLLPFPALH"	651. .7661	/gene="VC2544"	651. .7661	/gene="VC2544"	/note="similar to SP:P09200 GB:X12545 PID:41416 PID:537074 GB:U00096; identified by sequence similarity; putative"	/codon_start=1	/transl_table=1	/product="fructose-1,6-bisphosphatase"	/protein_id="AA95685.1"	/db_xref="GI:9657128"	/translation="MOKMRTLGEFIVKOHDFPHASGELSILASTIRLAKIVYREINIKAGIDITAGSNDNIOGEEFOOKDITVANKERFALALRQCVASAEDEAVAFSKELNNKAYVYLMPLDSSNIDVAVSGTIFSTIRRPSPGTPPTQEDFIQPGKQVLAAGVITGSSTMLVTTGNGVNGFTYDLSGFTFYSHNMRIPEMGKLYTSINEGATIRFPTGKRYKIRFCQENVPEEGRPYTSRYIGSLVADHRNLKGGITLYLPSYTSHPNGKLRLLYCNPMAPFLIEQAGGLASDGARRIMDIKPTLELHQRPPEFVSGKNMVHKEFLETYPD"	7831 . .8361	/gene="VC2545"	7831 . .8361	/gene="VC2545"	/note="similar to GB:U14003 SP:P17288 PID:450373 PID:537068 GB:U00096; identified by sequence similarity; putative"	/codon_start=1	/transl_table=1	/product="inorganic pyrophosphatase"	1.24e-18	Length: 14390	Matches: 164	Conservative: 104	Mismatches: 238	Indels: 226	Gaps: 20
gene	6334 . .6597	/gene="VC2543"	6334 . .6597	/note="identified by Glimmer2; putative"	/codon_start=1	/transl_table=1	/product="hypothetical protein"	/protein_id="AA95684.1"	/db_xref="GI:9657127"	/translation="MNVHKNPALFLCDWNGSIIPIRYAKASAAVQGVNSKMSITIKVRSPSVSITPVKRSKRYNSLQTRCIAQSIPSEDPITLLPFPALH"	651. .7661	/gene="VC2544"	651. .7661	/gene="VC2544"	/note="similar to SP:P09200 GB:X12545 PID:41416 PID:537074 GB:U00096; identified by sequence similarity; putative"	/codon_start=1	/transl_table=1	/product="fructose-1,6-bisphosphatase"	/protein_id="AA95685.1"	/db_xref="GI:9657128"	/translation="MOKMRTLGEFIVKOHDFPHASGELSILASTIRLAKIVYREINIKAGIDITAGSNDNIOGEEFOOKDITVANKERFALALRQCVASAEDEAVAFSKELNNKAYVYLMPLDSSNIDVAVSGTIFSTIRRPSPGTPPTQEDFIQPGKQVLAAGVITGSSTMLVTTGNGVNGFTYDLSGFTFYSHNMRIPEMGKLYTSINEGATIRFPTGKRYKIRFCQENVPEEGRPYTSRYIGSLVADHRNLKGGITLYLPSYTSHPNGKLRLLYCNPMAPFLIEQAGGLASDGARRIMDIKPTLELHQRPPEFVSGKNMVHKEFLETYPD"	7831 . .8361	/gene="VC2545"	7831 . .8361	/gene="VC2545"	/note="similar to GB:U14003 SP:P17288 PID:450373 PID:537068 GB:U00096; identified by sequence similarity; putative"	/codon_start=1	/transl_table=1	/product="inorganic pyrophosphatase"	1.24e-18	Length: 14390	Matches: 164	Conservative: 104	Mismatches: 238	Indels: 226	Gaps: 20
gene	6334 . .6597	/gene="VC2543"	6334 . .6597	/note="identified by Glimmer2; putative"	/codon_start=1	/transl_table=1	/product="hypothetical protein"	/protein_id="AA95684.1"	/db_xref="GI:9657127"	/translation="MNVHKNPALFLCDWNGSIIPIRYAKASAAVQGVNSKMSITIKVRSPSVSITPVKRSKRYNSLQTRCIAQSIPSEDPITLLPFPALH"	651. .7661	/gene="VC2544"	651. .7661	/gene="VC2544"	/note="similar to SP:P09200 GB:X12545 PID:41416 PID:537074 GB:U00096; identified by sequence similarity; putative"	/codon_start=1	/transl_table=1	/product="fructose-1,6-bisphosphatase"	/protein_id="AA95685.1"	/db_xref="GI:9657128"	/translation="MOKMRTLGEFIVKOHDFPHASGELSILASTIRLAKIVYREINIKAGIDITAGSNDNIOGEEFOOKDITVANKERFALALRQCVASAEDEAVAFSKELNNKAYVYLMPLDSSNIDVAVSGTIFSTIRRPSPGTPPTQEDFIQPGKQVLAAGVITGSSTMLVTTGNGVNGFTYDLSGFTFYSHNMRIPEMGKLYTSINEGATIRFPTGKRYKIRFCQENVPEEGRPYTSRYIGSLVADHRNLKGGITLYLPSYTSHPNGKLRLLYCNPMAPFLIEQAGGLASDGARRIMDIKPTLELHQRPPEFVSGKNMVHKEFLETYPD"	7831 . .8361	/gene="VC2545"	7831 . .8361	/gene="VC2545"	/note="similar to GB:U14003 SP:P17288 PID:450373 PID:537068 GB:U00096; identified by sequence similarity; putative"	/codon_start=1	/transl_table=1	/product="inorganic pyrophosphatase"	1.24e-18	Length: 14390	Matches: 164	Conservative: 104	Mismatches: 238	Indels: 226	Gaps: 20
gene	6334 . .6597	/gene="VC2543"	6334 . .6597	/note="identified by Glimmer2; putative"	/codon_start=1	/transl_table=1	/product="hypothetical protein"	/protein_id="AA95684.1"	/db_xref="GI:9657127"	/translation="MNVHKNPALFLCDWNGSIIPIRYAKASAAVQGVNSKMSITIKVRSPSVSITPVKRSKRYNSLQTRCIAQSIPSEDPITLLPFPALH"	651. .7661	/gene="VC2544"	651. .7661	/gene="VC2544"	/note="similar to SP:P09200 GB:X12545 PID:41416 PID:537074 GB:U00096; identified by sequence similarity; putative"	/codon_start=1	/transl_table=1	/product="fructose-1,6-bisphosphatase"	/protein_id="AA95685.1"	/db_xref="GI:9657128"	/translation="MOKMRTLGEFIVKOHDFPHASGELSILASTIRLAKIVYREINIKAGIDITAGSNDNIOGEEFOOKDITVANKERFALALRQCVASAEDEAVAFSKELNNKAYVYLMPLDSSNIDVAVSGTIFSTIRRPSPGTPPTQEDFIQPGKQVLAAGVITGSSTMLVTTGNGVNGFTYDLSGFTFYSHNMRIPEMGKLYTSINEGATIRFPTGKRYKIRFCQENVPEEGRPYTSRYIGSLVADHRNLKGGITLYLPSYTSHPNGKLRLLYCNPMAPFLIEQAGGLASDGARRIMDIKPTLELHQRPPEFVSGKNMVHKEFLETYPD"	7831 . .8361	/gene="VC2545"	7831 . .8361	/gene="VC2545"	/note="similar to GB:U14003 SP:P17288 PID:450373 PID:537068 GB:U00096; identified by sequence similarity; putative"	/codon_start=1	/transl_table=1	/product="inorganic pyrophosphatase"	1.24e-18	Length: 14390	Matches: 164	Conservative: 104	Mismatches: 238	Indels: 226	Gaps: 20
gene	6334 . .6597	/gene="VC2543"	6334 . .6597	/note="identified by Glimmer2; putative"	/codon_start=1	/transl_table=1	/product="hypothetical protein"	/protein_id="AA95684.1"	/db_xref="GI:9657127"	/translation="MNVHKNPALFLCDWNGSIIPIRYAKASAAVQGVNSKMSITIKVRSPSVSIT																										

Oy	292	methrthrvAlAlaAspGluValProLeuLeuIleGlyAspValPheHisGlyLys	311
Db	13904	TTTTCAACGTTTGAATCCGCCCAAGTCGTTTGAGAGTCGATGCGCGCTTAACCAAGTGC	1384
Oy	312	TyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGly	331
Db	13844	TATGCAACCTTCMAATCCGGCATTCGTAACCTTGGCGTTGCCAAAGAGCTACTTAATGCT	1378
Oy	332	ArgTrpLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeu	351
Db	13784	GATTTTGACGCCACCGCGCTGGAAGTCATGTCACAGCGTAATCAAGCTCGCGGTATTTTG	1372
Oy	352	IleTyrAspThrGlyThrGluTyrArgPheAspGluValValPhePheThrIleAspPro	371
Db	13724	CACTTTGATACGGGATCCGTTACTGTTT-----	1369
Oy	372	LysThrAsnGluLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeuGlu	391
Db	13694	-----GCAGCGACACGCGTCGAGCGCAAGTCAAATCGATGAAATCCGCGATG-----	1364
Oy	392	GlnLeuLeuThrValAsnMetGlyGluAlaTyrAlaSerLeuGlnAlaValAlaArgAlaLeuSer	411
Db	13646	TCCGTGGCCCTTTCAAAACAAGCGGACCTTATTTGGTCGCAAGTCGCGCAATTAAAC	1358
Oy	412	AsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGlu	431
Db	13586	CAAAATCTTTCCAATACAGATTGGTCTCTCGTGCTTTGTTCAG-----CCGCAAC	1353
Oy	432	ArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerSerArgThrGlu	451
Db	13535	TTTTCCGACATTG-----	1352
Oy	452	ProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAsp	471
Db	13524	-----	1352
Oy	472	GlyIleLeuMetAspLysSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLys	491
Db	13524	-----	1352
Oy	492	LeuAsnLeuValAlaAlaLysAlaArgHisLysLeuTyrAspMetProAspArgValLeu	511
Db	13524	-----	1352
Oy	512	AlaIleAsnHisAspAspGlyValAsnArgSerTleLeuGlyArgTlleSerAspAlaVal	531
Db	13523	-----CATATGGGG-----	1351
Oy	532	SerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuPro	551
Db	13515	-----	1351
Oy	552	GluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysVal	571
Db	13514	-----CGTGAGTTG	13506
Oy	572	ProLeuTyrValPheValAlaSerAspLysProAlaArgAspGlyGlnIleGlyLeuGlyTrp	591
Db	13505	CCGATCAAGCTACCTTACCAACCGCAACGCAACCAATCAACTGAAACCGGTTAGGCTAT	13468
Oy	592	GlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArgAsp	611
Db	13445	TCCACGACGCTTGGAGTACGCTGTCATTAATAATGMAAACAACCTTGGGTGAACGCCA	13386
Oy	612	GlyTyrGlnAlaGlyAlaGluLeuAspGluSerGluAspLysLysGlyValLysLeuTyr	631
Db	13385	GGGCAACAGTTTGAACAGCACTTCTCTCTTCTATACAGACAGACATTCACGGCGGT	13326
Oy	632	AlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyLysGln	651
Db	13325	TACAAAGATCCCACTGGAAAGATCCGTTAAATGAGTACTAGTACGATTACAGTACCGGATGAAG	13266

QY	652	GlnGluValAlaPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHis	671
Db	13365	-----CATTTGGACAAAG-----CGTGAATCCGACGCTTGGAATCT	13230
QY	672	GluIleSer-----ArgSerIleIleGlnAsnGlyGlyTrpAsnArgThrTyrSerLeu	689
Db	13229	AATGATACACCTGTAACGCCAATGGCAATTGGATGGATGGCGGCTGGCATGCCAGGCTTTTAAT	13170
QY	690	ArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGlnIleThrTrpGlnAspLeuPro	709
Db	13169	CGTTATCTCTTACGGAATATGCT-----CAGGGTTTGACA	13134
QY	710	ValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHis	729
Db	13133	GATGAT-----MACGCCAATTTTATTGGCGGGGAGACTTATACC	13092
QY	730	LysThr---ValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSer	748
Db	13091	CGTACGGCGGACCTCGCTCCACCATGGCTTACTGACTTGGGCGCATTAACAGACGATTACC	13032
QY	749	LeuGlnValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaAlaGly	768
Db	13031	TTGGAAATATGGCCATCCAGCATTCCTTCGAAACACGGGTATTCGTTTACAAACCGGA	12972
QY	769	IleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMet	788
Db	12971	TCCTTGTGGTTACGACTTATGCGCCAAAT-----CACCGCGCC	12933
QY	789	ThrClyGlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArg	808
Db	12932	TTAGTTCCGCTTGATGTGGAGCCAACTTGCTGATGAATTTGACCAACACTTCGCCCTCG	12873
QY	809	LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisPheSerLeuSer	828
Db	12872	CTGCGTTTCTTGCCGCTGTGTATACAAATTTGCGCGGATACGCTTACAGTCGATTTTCA	12813
QY	829	ProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGlu	848
Db	12812	CCGCAAGATGCCAGTGGCGCTTTGACAGCGGCGGAATATCATGCGCCACAGCTCAATCGAA	12753
QY	849	TyrAsnTyrGlyIlePheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAla	868
Db	12752	TATCATACCGCTTAAACCGCAACGTGTGGCGGCATTTGATGATGTTCGTGTGTGG	12693
QY	869	TyrAspLysGlyPheThrAsnAspThr-----LysIleGlyAlaGlyValGlyVal	885
Db	12692	TTT-----AACCATATCTCGATGAGAAAGAGCGTCTGCTGATCGGATC	12648
QY	886	ArgTyrPalaSerProValGlyGlnValArgValAspValAlaThrGlyVal---LysGlu	904
Db	12647	CGCTGATTTTGGCCCGTGTGGCCCTATTCGGCTCGATTTTCGCATGTGGGCGCTGGATCCGCG	12588
QY	905	GluGlyAsnProIleLysLeuHisPhePheIleGly	916
Db	12587	CCGGGCGATGAGCTCAAAATCCACTTCACTTAGG	12552
LOCUS	AX033468	1737 bp	DNA
DEFINITION	Sequence 1 from Patent WO0047737.	linear	PAT 21-SEP-2000
ACCESSION	AX033468		
VERSION	AX033468.1	GI:10280229	
KEYWORDS			
SOURCE	Haemophilus influenzae.		
ORGANISM	Haemophilus influenzae Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Haemophilus.		
REFERENCE	Ruelle J-L, and Thonnard J.		
AUTHORS	Haemophilus influenzae rd outer membrane sequences used as vaccine		
TITLE	Patent: WO 0047737-A 1 17-AUG-2000;		
JOURNAL	RUELLE JEAN LOUIS (BE) ; SMITHKLINE BEECHAM BIOLOG (BE) ; THONNARD JOELLE (BE)		

[illegible]

gene	4261. .4983	/gene="H10694"	/note="similar to GB:U00096 SP:P75966 PID:1787380 percent identity: 66.29; identified by sequence similarity; putative"	
gene	4261. .4983	/transl_table=1	/codon_start=1	
gene	4261. .4983	/product="conserved hypothetical protein"	/protein_id="AAC22354.1"	
gene	4261. .4983	/db_xref="GI:1573697"		
gene	4261. .4983	/translation="YMLQAFQNRNHHKLKKNFSAQGLPSKGSANVFNRHFRPKLPKRLTSLSDKLVYLFNMPQAFVLTQFTDEQSKLTKDLSIPNYAAQRDRDSGLILTNNGEHLADLPKDEKTEKTYWQVGEIPEEDLQARKGVELKDGVTSAKARLISEPNLMEENPIRREKRNIPTSWLEIKISEGRNRQVRRMTAHGPTLLVRSVNGLLSINGLENGSEFRILSDIAOLFQTVKL"		
gene	4261. .4983	complement(5109. .6080)		
gene	4261. .4983	/gene="H10695"		
gene	4261. .4983	complement(5109. .6080)		
gene	4261. .4983	/note="similar to GB:U06129 SP:P29014 PID:147343 GB:U00096 PID:1788848 percent identity: 53.62; identified by sequence similarity; putative"		
gene	4261. .4983	/codon_start=1		
gene	4261. .4983	/transl_table=1		
gene	4261. .4983	/product="exopolyphosphatase, putative"		
gene	4261. .4983	/protein_id="AAC22355.1"		
gene	4261. .4983	/db_xref="GI:1573698"		
gene	4261. .4983	/translation="MNSDILEPKHRGNVREITAIIDLGSNFHMTIARIANGSIOVLRKQVLAEGIDENAVLNQEAITRGVNCALFARLQGFPMENNVAGTTLBRVAVNDEFLROAKVPEPPIINIIISGOTEAKTICAGCHQPERKGVITIDIGGSEMIIGDDFTPLMAESRHMGCVSPATQFTDITISEPNQRKQSAVKIEDLGLETRKLGQSVLGSSTITATQVAVIATNLDPNSTITAEIRNALIEOTLDKHFTEININGLQDRVDFVPGALISAVFDVPIHQMRSDGALRGVYISLEKNFQVADIRASTA"		
gene	4261. .4983	complement(6088. .9984)		
gene	4261. .4983	/gene="H10696"		
gene	4261. .4983	complement(6088. .9984)		
gene	4261. .4983	/note="similar to GB:U14003 SP:P39322 PID:537063 GB:U00096 percent identity: 34.87; identified by sequence similarity; putative"		
gene	4261. .4983	/codon_start=1		
gene	4261. .4983	/transl_table=1		
gene	4261. .4983	/product="conserved hypothetical protein"		
gene	4261. .4983	/protein_id="AAC22356.1"		
gene	4261. .4983	/db_xref="GI:1573699"		
gene	4261. .4983	/translation="MTBQIOPSETPSKSPKPNKKHWRKAVCIGSAYVILPYLVAGALISDQACKSLIQLVDMKIDSFSEVQEGQONLVLKNVYQAGIEHTIAQRLDODDGLGSEVREVLRFDTLNKPTIATNTALPLPSADNKGSGMKISLISINAMLVNODLSVNIIDQTSITLGNFKSANSYLNNEKGTITAPPEINDISYIAKISLSEVSEKARDPNKPVDMAIIBOSLTPARLGNSEILLPFDLHPIEISCKNMQYQVNEKGETLOSVEMSLIIASVDNLDQLOKLAVESLSGNLSQKLOLDGMDPLDNLKSHLEPKDQKEILPISVDVLTLSGLKSTALSTLKTGVLAELNGVNOADQKMPMLTLNVAKQYTTLVNMTPLKINDVTLKLTGDLNHAELKGVAGAGNYPASQVELNADGKLTVEYVARKLGIDLSQKSEFVGNAMNKCANNMDIOADLEKNMIAFVPMAPATLSGLKHSRFACSSQGOVEYEVADLNGMLSAKPTLSLGSKATONVYLTLPDLOIKGNYILKASGVLDLHSDPDLINAPLRLGLMSLDIGRKVGAISIOVLTPLDIDTISNHLIGPOLAKSIKGINNASSLSGKLNIKARLHYGNIKLHLLDDLSQDBONKILIKSGEYVAAANLQIGHGDRTLQONKGTISOYKEFTPIGDKVSNQALAVSDNKOQNTASHCWQNTDVELCEPOAFNAGKQGNIPFOKRVNLDLVNKLIEONSLKQVQGVNVAWTFDIPQFLTANDVGNHLAFSOKLDYRTFKLYIPKLTLANDIQNNMLVTKTDIVHNQGRVIGDPIHLNDLAKRQIGTTLAIERLNLSIANQLTSGSVGEVYAKLSFEGNLEKPLINDFNINRITRKLKSPVNTITDGDIALRFNDNRSTLOGKIKTYDSHLNIGRANMANIEHWTFLMAQNNPNDIPSMARKLRSPNTITRANKPELMSGVDTIPARIKTDSLPORTAPVYSEDEVILLNGPHKSKSEELIKREFAKTSCHEIRSDLWNTGKDYSLAYGLKTNDELTVSADQGNGLDFOQINDPVTGRVASFODDILIRKGLISFSQQAQPLUNLAINPPELTMSDKRTIAGVAVGVIADSPETVITRSEPSKPDQALSVLLGRLESSGEVASTSVGAVALIIGLISKSGKLTGVSIGEVFGIOLMLNGISGVADKSKVYSGNITNRLQIKYGVLF		
gene	4261. .4983	complement(6088. .9984)		
gene	4261. .4983	/gene="H10696"		
gene	4261. .4983	complement(6088. .9984)		
gene	4261. .4983	/note="similar to GB:U14003 SP:P39322 PID:537063 GB:U00096 percent identity: 34.87; identified by sequence similarity; putative"		
gene	4261. .4983	/codon_start=1		
gene	4261. .4983	/transl_table=1		
gene	4261. .4983	/product="conserved hypothetical protein"		
gene	4261. .4983	/protein_id="AAC22356.1"		
gene	4261. .4983	/db_xref="GI:1573699"		
gene	4261. .4983	/translation="MTBQIOPSETPSKSPKPNKKHWRKAVCIGSAYVILPYLVAGALISDQACKSLIQLVDMKIDSFSEVQEGQONLVLKNVYQAGIEHTIAQRLDODDGLGSEVREVLRFDTLNKPTIATNTALPLPSADNKGSGMKISLISINAMLVNODLSVNIIDQTSITLGNFKSANSYLNNEKGTITAPPEINDISYIAKISLSEVSEKARDPNKPVDMAIIBOSLTPARLGNSEILLPFDLHPIEISCKNMQYQVNEKGETLOSVEMSLIIASVDNLDQLOKLAVESLSGNLSQKLOLDGMDPLDNLKSHLEPKDQKEILPISVDVLTLSGLKSTALSTLKTGVLAELNGVNOADQKMPMLTLNVAKQYTTLVNMTPLKINDVTLKLTGDLNHAELKGVAGAGNYPASQVELNADGKLTVEYVARKLGIDLSQKSEFVGNAMNKCANNMDIOADLEKNMIAFVPMAPATLSGLKHSRFACSSQGOVEYEVADLNGMLSAKPTLSLGSKATONVYLTLPDLOIKGNYILKASGVLDLHSDPDLINAPLRLGLMSLDIGRKVGAISIOVLTPLDIDTISNHLIGPOLAKSIKGINNASSLSGKLNIKARLHYGNIKLHLLDDLSQDBONKILIKSGEYVAAANLQIGHGDRTLQONKGTISOYKEFTPIGDKVSNQALAVSDNKOQNTASHCWQNTDVELCEPOAFNAGKQGNIPFOKRVNLDLVNKLIEONSLKQVQGVNVAWTFDIPQFLTANDVGNHLAFSOKLDYRTFKLYIPKLTLANDIQNNMLVTKTDIVHNQGRVIGDPIHLNDLAKRQIGTTLAIERLNLSIANQLTSGSVGEVYAKLSFEGNLEKPLINDFNINRITRKLKSPVNTITDGDIALRFNDNRSTLOGKIKTYDSHLNIGRANMANIEHWTFLMAQNNPNDIPSMARKLRSPNTITRANKPELMSGVDTIPARIKTDSLPORTAPVYSEDEVILLNGPHKSKSEELIKREFAKTSCHEIRSDLWNTGKDYSLAYGLKTNDELTVSADQGNGLDFOQINDPVTGRVASFODDILIRKGLISFSQQAQPLUNLAINPPELTMSDKRTIAGVAVGVIADSPETVITRSEPSKPDQALSVLLGRLESSGEVASTSVGAVALIIGLISKSGKLTGVSIGEVFGIOLMLNGISGVADKSKVYSGNITNRLQIKYGVLF		
gene	4261. .4983	complement(6088. .9984)		
gene	4261. .4983	/gene="H10696"		
gene	4261. .4983	complement(6088. .9984)		
gene	4261. .4983	/note="similar to GB:U14003 SP:P39322 PID:537063 GB:U00096 percent identity: 34.87; identified by sequence similarity; putative"		
gene	4261. .4983	/codon_start=1		
gene	4261. .4983	/transl_table=1		
gene	4261. .4983	/product="conserved hypothetical protein"		
gene	4261. .4983	/protein_id="AAC22356.1"		
gene	4261. .4983	/db_xref="GI:1573699"		
gene	4261. .4983	/translation="MTBQIOPSETPSKSPKPNKKHWRKAVCIGSAYVILPYLVAGALISDQACKSLIQLVDMKIDSFSEVQEGQONLVLKNVYQAGIEHTIAQRLDODDGLGSEVREVLRFDTLNKPTIATNTALPLPSADNKGSGMKISLISINAMLVNODLSVNIIDQTSITLGN		

[illegible]

Oy	580	AspLysProArgAspGlyInIleGlyLeuGlyTyrGlySerAspThrCysIleThrArgLeu	599
Db	10944	CCTAAAAAAATGCCGATGCACTCGGTGGCCCTTTCTACTGATGCGCGCTTCACGGA	10885
Oy	600	ValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaLeu	619
Db	10884	CAAAATAGCCTGGCAAAACCTGGATTATTAATACCAGCCGACATAGTTGGCTCAATCTT	10825
Oy	620	ArgLeuSerGlnAspLysGlyGlyAlaLys----LeuTyrAlaIleThrLysProLeuSerHis	638
Db	10824	TATCTCTCTGCACCAAAACAACTCTACGAGCACTTATTCGATGATCCACTGCTTAAAAAT	10765
Oy	639	ProLeuAsnAspGlnLeuArgAlaIleThrLeuGlyTyrGlnGlnGlnValPheGlyHisSer	658
Db	10764	CCATTAAATTTATTACTATGATTTTGGCCGCGTGGGAA-----GGGGAANA	10717
Oy	659	ThrAsnGlyPheAspLeuSerThrArgThrIleLeuGluHisGluIleSerArgSerIleIle	678
Db	10716	GAGAAAT-----GATACCAATCGAGAGACTGACTGCTTCTCAGCGTTACCTATTGGAA	10663
Oy	679	GlnAsnGlyGlyTyrPasnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThr	698
Db	10662	AATGCCCATGCTGGGCAATATTTTGGCGACTTCTGATGCGATAGACAGTTT---ACA	10606
Oy	699	GlnAlaProProGluIleThrTyrPcinAspLeuProValAspPheValAsnGlyLysProSer	718
Db	10605	CAACGGCAAT-----	10597
Oy	719	GlnGluAlaLeuLeuAlaGlyValAlaValaHisLysThrValAlaAspAsnValaAsn	738
Db	10596	-----ATCACTGATAAACCTTA-----CTTCTTTAT	10570
Oy	739	ProMetArgGlyTyrArg-----	744
Db	10569	CCAACTGTGGATTACTGCTACTGATTAACGCTGGTGTCTTCCACTGGGCGCAT	10510
Oy	745	---GlnArgTyrSerLeuGlnValaGlySerSerGlyLeuValSerAspAlaSerMetAla	763
Db	10509	GTGCAAAAAATTACTTTTGATTTTAAGCAACGAATTTGGCTATCGAAATCTTCTTTATA	10450
Oy	764	IleAlaArgAlaGlyIleSerGlyValaTyrSerPheGlyAspAsnAlaTyrGlySerAsn	783
Db	10449	AAAGTCAACGCACTACGCGCGGTCTGTAATGACGAANAAT-----	10405
Oy	784	ArgAlaHisGlnMetThrClyGlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsn	803
Db	10404	-----CATCGTGTGCTGCTGCTGCTCAAAATCGGCTATTATTCATCAAAACGATTTGAA	10351
Oy	804	HisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAla	823
Db	10350	AAATTTCCGCCCTACACTGCGTTCTTCTGCTGGCGCATCGTAGTGTCGCGCGTATACGC	10291
Oy	824	HisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlnValLeuAla	843
Db	10290	TATAAAAAAATTGCCCTTAAAAAATGAATGGAATTTGCTGGCGTTCGCGATTGCTT	10231
Oy	844	ValGlyThrAlaGluTyrAsnProTyrCyluPheMetLysAspLeuArgLeuAlaValPheGly	863
Db	10230	ACCACCTCTTTAGATATATCAATATCACTTATCCAAATTTGGTGGCGGCACTTTTGCA	10171
Oy	864	AspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyAla	883
Db	10170	GATAGTGAATTAGTCGGCGATATATTAACACAGCAAAAGACGTCCGTTATGGCAGAGCGTT	10111
Oy	884	GlyValaArgTTPalaSerProValGlyGlnValaArgValaAspValaIleThrGlyValaLys	903
Db	10110	GGTGCGCTTGGCGCATCCCGAGTTGGTCCGATTAATTTGATATTCGACACACCCATTCTGT	10051
Oy	904	GluGlnGlyAsnProIleLysLeuHisPheIleGly	916
Db	10050	CATAAAGCTAACACGCAAAATATTCATTTTACATCGGA	10012

LOCUS	243184 bp	DNA	linear	BCT 07-MAR-2001
DEFINITION	Escherichia coli O157:H7 DNA, complete genome, section 19/20.			
ACCESSION	AP002568 BA000007			
VERSION	AP002568.1 GI:13364484			
KEYWORDS	Escherichia coli O157:H7 (strain:O157:H7, sub_strain:RIMD 0509952) DNA.			
ORGANISM	Escherichia coli O157:H7 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.			
REFERENCE	1 (sites)			
AUTHORS	Makino, K., Yokoyama, K., Kubota, Y., Yutsudo, C.H., Kimura, S., Kurokawa, K., Ishii, K., Hattori, M., Tatsuno, I., Abe, H., Iida, T., Yamamoto, K., Ohnishi, M., Hayashi, T., Yasunaga, T., Honda, T., Sasakawa, C. and Shlnagawa, H.			
TITLE	Complete nucleotide sequence of the prophage VT2-Sakai carrying the verotoxin 2 genes of the enterohemorrhagic Escherichia coli O157:H7 derived from the Sakai outbreak			
JOURNAL	Genes Genet. Syst. 74 (5), 227-239 (1999)			
MEDLINE	20198780			
REFERENCE	2 (sites)			
AUTHORS	Ohnishi, M., Murata, T., Nakayama, K., Kuhara, S., Hattori, M., Kurokawa, K., Yasunaga, T., Yokoyama, K., Makino, K., Shlnagawa, H. and Hayashi, T.			
TITLE	Comparative analysis of the whole set of rRNA operons between an enterohemorrhagic Escherichia coli O157:H7 Sakai strain and an Escherichia coli K-12 strain MG1655			
JOURNAL	Syst. Appl. Microbiol. 23 (3), 315-324 (2000)			
MEDLINE	20557356			
REFERENCE	3 (sites)			
AUTHORS	Yokoyama, K., Makino, K., Kubota, Y., Watanabe, M., Kimura, S., Yutsudo, C.H., Kurokawa, K., Ishii, K., Hattori, M., Abe, H., Iida, T., Yamamoto, K., Hayashi, T., Yasunaga, T., Honda, T., Sasakawa, C. and Shlnagawa, H.			
TITLE	Complete nucleotide sequence of the prophage VT1-Sakai carrying the Shiga toxin 1 genes of the enterohemorrhagic Escherichia coli O157:H7 strain derived from the Sakai outbreak			
JOURNAL	Gene 258 (1-2), 127-139 (2000)			
MEDLINE	20564182			
REFERENCE	4 (sites)			
AUTHORS	Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K., Yokoyama, K., Han, C.-G., Ohtsubo, E., Nakayama, K., Murata, T., Tanaka, M., Tobe, T., Iida, T., Takami, H., Honda, T., Sasakawa, C., Ogasawara, N., Yasunaga, T., Kuhara, S., Shiba, T., Hattori, M. and Shlnagawa, H.			
TITLE	Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12			
JOURNAL	DNA Res. 8 (1), 11-22 (2001)			
MEDLINE	21156231			
REFERENCE	5 (bases 1 to 243184)			
AUTHORS	Ohnishi, M., Kurokawa, K., Makino, K., Yasunaga, T., Shlnagawa, H. and Hayashi, T.			
TITLE	Direct Submission			
JOURNAL	Submitted (26-JUN-2000) Ken Kurokawa, Osaka University, Genome Information Research Center, 3-1, Yamadaoka, Suita, Osaka 565-0871, Japan (E-mail:ken@gen-info.osaka-u.ac.jp, URL:http://www.gen-info.osaka-u.ac.jp/, Tel:81-6-6879-8365, Fax:81-6-6879-2047)			
MEDLINE	21156231			
REFERENCE	genome project.			
COMMENT	Location/Qualifiers			
FEATURES	1. 243184			
SOURCE	/organism="Escherichia coli O157:H7"			
	/strain="O157:H7"			
	/sub_strain="RIMD 0509952"			
	/db_xref="taxon:83334"			
	466.. 2115			
gene	/gene="ECs5008"			
	466.. 2115			
gene	/gene="ECs5008"			
	/note="similar to PGL-ECOLI g11790457 percent identity 100 in 549 aa (Conserved in E.coli K-12)"			
CDS	/codon_start=1			

/evidence-not_experimental
 /transl_table=11
 /product="glucose-6-phosphate isomerase"
 /protein_id="BAB38431.1"
 /db_xref="GI:13364485"
 /translation="MKNIPTOTAAQALOKHFDEMDKVTIADLFKADGRFSFSAW
 FDDMLVDYSKRRIETELAKLODLAKEDCLAGATISMSEFGKIRNRENKAVHVALR
 NRSNPILVGDGDWMEVNAVLKEMTSPSALITSGCKGKGTGCAIDPVNMGISGSD
 GPWYTEALRPKKNHMHVSVSDVSTHIAEVLKKNPEETLLVLVASKITFTTQETMKN
 AHASDWPLKKAAGDEHNVKHFALASTNKAAGEFGIDVAMFEEMWVGGRYSLMSNA
 IGLSVLSGTFNPELLSGAHMDKHFTTPAEKMLPVLLALIGIYNFNGAETBA
 ILTPDYMHSEFAFPAQOQMSNGKTYDRNGNVVDYOTGPIINGEPTNGOAHAYOLIL
 HOGTWMPDEDFAPATITNHPDLSPHOKILSNFPAQTEALAFGKSRVBEQYRQGGKD
 PATIDVYVPEFEGCKRRPNSTILSRITPSGLCALILAEHKITFGOSVLIINIFFDOM
 GVELCKOLANRLIPELKDKEISHDSSINGLINRYKAMRG"
 2628. .2870
 /gene="ECS5009"
 2628. .2870
 /gene="ECS5009"
 /note="similar to yJbE_ECOLI g11790458 percent identity
 100 in 80 aa (conserved in E.coli K-12)."
 /codon_start=1
 /evidence-not_experimental
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAB38432.1"
 /db_xref="GI:13364486"
 /translation="MKKVLYGIFALATSAMAAPVOYGEAASNAATSVSAGSSAA
 TSVSTVSSAVGVALAATGGDGSNTGTTTTTSTQ"
 2954. .3622
 /gene="ECS5010"
 2954. .3622
 /gene="ECS5010"
 /note="similar to yJbF_ECOLI g11790459 percent identity
 96 in 222 aa (conserved in E.coli K-12)."
 /codon_start=1
 /evidence-not_experimental
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAB38433.1"
 /db_xref="GI:13364487"
 /translation="WVILPMLRVYKRPATILICLLQACSAATTKELGNSLMDISFCTP
 PGVOLTADIDOMPRASQWQMLNGGQLFVYLAFAEDGQKRWYTDODATITVYOHGRVLY
 KTLGSGDNLIEVNNLAADPLIPAOITIDGATWMTKWTGWEYQDOVRVATARSVEWGDST
 GTVKGSDETAARVLADEEYSTDOARWNNRYWIDSGQIROSEQYLGADYPRVYKTLIKI
 AAKQ"
 3619. .4356
 /gene="ECS5011"
 3619. .4356
 /gene="ECS5011"
 /note="similar to yJbG_ECOLI g11790460 percent identity
 97 in 245 aa (conserved in E.coli K-12)."
 /codon_start=1
 /evidence-not_experimental
 /transl_table=11
 /product="hypothetical protein"
 /protein_id="BAB38434.1"
 /db_xref="GI:13364488"
 /translation="MIKQITVILSLVGSASVFAAGTVFSGNSSEAKTLTGAEHLIIL
 DLVQGPRLANSWMPKQVISEELATAALEOQOQALITRLAEQGDSSADDAALINALQ
 QIQALKVTRGKINIDPDIVRAERANPLQGYITLWVGPPSTVILFGLISRGNDP
 FTPGQDVASVYSGSLGCGADSYAMVYPDGTOKAPAYAMNKRHVEPMPGSIYVG
 LADSWSEMPDALNNDIITLQRIQ"
 4356. .6452
 /gene="ECS5012"
 4356. .6452
 /gene="ECS5012"
 /note="similar to yJbH_ECOLI g11790461 percent identity
 99 in 698 aa (conserved in E.coli K-12)."
 /codon_start=1
 /evidence-not_experimental
 /transl_table=11
 /product="hypothetical protein"

[illegible]

Alignment Scores:			
Pred. No.:	8,77e-16	Length:	243184
Score:	434.50	Matches:	164
Percent Similarity:	36.92%	Conservative:	100
Best Local Similarity:	22.94%	Mismatches:	262
Query Match:	9.19%	Indels:	189
DB:	1	Gaps:	18

US-09-914-168-2 (1-919) x AP002568 (1-243184)

QY	211	ASnlelysaAlaaleuGlusplleThrcInluserAlametasplLeuansgLYser	230
			:::
Db	211421	AACGTCGTGGCGACGCTTTCACGATTGAAGAAGCATATAAGTGAGCCGACGCGCTTT	211460
QY	231	lIleProArglLeuArGInlThrAlaLeuValAlaAlaArGAlaValAlGlyTyr-----	248
			:::
Db	211481	CGCGACCGTCGATGATGCTATCCGGAGAGGCTGTGAACGCGGGGCTTTTACCAAGCCG	211540

Oy 249 -----AspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAlaIleIle 266
Db 211541 ACCATTGAATTTGATCTCCGTCACCGCCAAAGAAAGGGCGAGATTGATGCGCAAA 211600
Oy 267 HisAspLeuGlyGluProValIleAspTyrArgAlaValGluValArgIleGlyGly 286
Db 211601 GTCACGCCAGCGGCGTGGTGAATTTGGCGCACCGATGGTGAATTTGGCGGCGCG 211660
Oy 287 AlaAspAspLeuAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal 306
Db 211661 CGGACCGGATAAAGACTATTGAAATTTGCTCGATCTACGCCCG---GCTATTGGCACCGTA 211717
Oy 307 PheHisHisGlyLysTyrGlyLeuThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHis 326
Db 211718 CTGAACACCGCGCATATGAAATTTCAAAAGTCTTAACGACGATGCGGTGCGCTAAA 211777
Oy 327 GlyTyrPheAspGlyArgTyrPheAspArgSerValAspValIleLeuProAspAsnThr 346
Db 211778 GGTATTTCGATAGCGAAATTTACCAAGCGCAGCGCATGCGCTGCGCTGCTATAA 211837
Oy 347 AlaAspValSerLeuIleTyrAspThrGlyThrGlyTyrArgPheAspGluValAlaPhe 366
Db 211838 GCGTCTCGCATATGATGATATTAACACTGCGCAACGCTTACCGCTTGGCATGTGACTTT 211897
Oy 367 PheThrIleAspProLysThrAsnGluLeuThrThrAspProAspLysLeuProValLys 386
Db 211898 -----GAAAGATCACAAATCCCG 211915
Oy 387 ArgGluLeuGluGluGluLeuLeuThrValAsnMetGlyGluAlaLysAsnLeuAla 406
Db 211916 CATGAATACCTCAAAATTTGGTCCGCTTAAAGAGCGGATGAGTACCAATTCGAAAGAT 211975
Oy 407 ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGlu 426
Db 211976 CTGGCAGAACTGACCGCGCATCTCTCTACCGCTGTTAACTCGGCG----- 212026
Oy 427 IleValPheProGluArgGluGluIleGlnAsnAspGluValSerPheGluInSerSer 446
Db 212027 GTGGTGGCTCCACAA-----TTGTATAAGCGCGC 212056
Oy 447 SerSerArgThrGluProAlaGluValAspGluSerThrLeuGluProValIleGluThr 466
Db 212057 GAAAGCAAACTATTACA----- 212074
Oy 467 ValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsn 486
Db 212075 -----TTTCACG---GCGGTG-----GTTTCGCGC----- 212095
Oy 487 LeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetPro 506
Db 212095 ----- 212095
Oy 507 AspaAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArg 526
Db 212095 ----- 212095
Oy 527 IleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu 546
Db 212096 -----CGAACAGAAACACT 212110
Oy 547 ValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr 566
Db 212111 ATC----- 212113
Oy 567 GlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGln 586
Db 212114 -----GAA 212116
Oy 587 IleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsn 606
Db 212117 ACCGGGCGTTCGTTACTTACGACGACGTCGGACCGCGCTGAAGCGACGCGAATAAACCG 212176

Oy 607 LeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLys 626
Db 212177 TGGATGAACCTCATACGCTTACACGTCTACACACGACAGTATTTCCGCGCGAAGACG 212236
Oy 627 GlyValLysLeu---TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArg 645
Db 212237 ACCCTGACCTTCAGCATAAATTCGCGCTCTGAAAGATCCACGTAATATTAATTCG 212296
Oy 646 AlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSer 665
Db 212297 GTCAGAGCGCGTTTAAAGCCACT-----GACCTGAACGATACCGAGCTTGAC 212344
Oy 666 ThrArgThrLeuGluHisGluLeuSerArgSerIleIleGlnAsnGlyTyrPasnArg 685
Db 212345 TCCACTACGCTG-----GTGGCTTCCTCGTACTGGGATCTCCACAGCGCTGGCACGCT 212398
Oy 686 ThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTyr 705
Db 212399 GCCATTAACTGCGCTGAGCTTCGAC----- 212425
Oy 706 GlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeu----- 723
Db 212426 -----CACTTACTCAGAGTGAATTAACCAACACGACGATGCTTTAT 212470
Oy 724 AlaGlyValAlaValHisLysThrValAlaAlaAspAsnLeuValAsnProMetArgGlyTyr 743
Db 212471 CCTGGGTGATGATTAAGCCGACGCGCTCTGTTGGCTGATGCCAACCTGGGCGAC 212530
Oy 744 ArgGluArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
Db 212531 TCCCAACGCTACTCTATGACTCTCAACACGAGCGCTGGGGGTTCAGATGTCGATTCTCC 212590
Oy 764 IleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsn 783
Db 212591 GTTTTCAGCGCGAGAACGCTGATCCGACACTGACATGATCGC----- 212635
Oy 784 ArgAlaHisGluMetThrGlyGlyIleGlnAlaGlyTyrIleThrPheSerAspAsnPheAsn 803
Db 212636 -----CAGCTGTTTGTGACACCGCGACGCTGGCGGATTAACACCGGTGATTTCCGAC 212689
Oy 804 HisValProTyrArgLeuArgPhePheAlaGlyIleAspInSerIleArgGlyTyrAla 823
Db 212690 AAGTACCGCGGATCTCGCTTCTTCGCGGGGCGCATGCGATTTCCGCGCTACAAA 212749
Oy 824 HisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAla 843
Db 212750 TACAAATCTATCCGCCGAAATACGCTAACGCTGACCTGAAGGGCGCTCGAAGTTGATA 212809
Oy 844 ValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGly 863
Db 212810 ACCGATCGCTGGATTAACGACTACACGTCGACCGGAAATGCTGGCGCGGTGTTGTC 212869
Oy 864 AspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyVal 883
Db 212870 GATAGTGGCGAAGCGGTAACGATATTCGCGCGACGACTTTTAAACGCGTACCGGGGTC 212929
Oy 884 GlyValArgTyrAlaSerProValGlyGlnValArgValAspValAlaThrGlyVal--- 902
Db 212930 GGGGTACGCTGGGAATCCGCGGTGGGCGCAATCAACTGATTTTGGCGTACCGGTGCGG 212989
Oy 903 ---LysGluGluGlyLysnProIleLysLeuHisPhePheIleGly 916
Db 212990 GATAAAGCAACACGCG-----TTACAGTTTATCATTCGCT 213025
RESULT 15
LOCUS AE000493 10819 bp DNA linear BCT 01-DEC-2000
DEFINITION Escherichia coli K12 MG1655 section 383 of 400 of the complete
genome.
ACCESSION AE000493 U00096
VERSION AE000493.1 GI:2367360
KEYWORDS Escherichia coli K12.

ORGANISM	Escherichia coli K12	CDS	156..656
REFERENCE	Bacteria: Proteobacteria; gamma subdivision: Enterobacteriaceae; Escherichia.		/gene="ytfI"
AUTHORS	1 (bases 1 to 10819)		/function="orf; unknown"
	Blattner,F.R., Plunkett,G. III, Bloch,C.A., Perna,N.T., Burland,V., Riley,M., Coliado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F., Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J., Mau,B. and Shao,Y.		/note="Q166"
TITLE	The complete genome sequence of Escherichia coli K-12		/codon_start=1
JOURNAL	Science 277 (5331), 1453-1474 (1997)		/transl_table=11
MDLINE	97426617		/product="orf; hypothetical protein"
REFERENCE	2 (bases 1 to 10819)	gene	/protein_id="AAC77172.1"
AUTHORS	Blattner,F.R.		/db_xref="GI:1790661"
TITLE	Direct Submission		/translation="MLPRI RHNNF IGAVEL FVKSS YTKHNNFN NIHAFK RWDI SNYSD LTLREFRCAT QIDKSGYVLSKNETVHAMDKFLISFLKNGAYETWTLR GSGFEIEEIPIT INEYNSFMDKRNREPLEQDDCPGIFYKKNLSUTYRKGLKVIETRLVTCI"
JOURNAL	Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.	CDS	complement(1156..1710)
	Email: ecollegenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459		/gene="ytfJ"
REFERENCE	3 (bases 1 to 10819)		/note="b4216"
AUTHORS	Blattner,F.R.		complement(1156..1710)
TITLE	Direct Submission		/gene="ytfJ"
JOURNAL	Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.		/function="orf; unknown"
	Email: ecollegenetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459		/note="f184; 100 pct identical amino acid sequence and equal length to YTFJ_ECOLI SW: P39318"
REFERENCE	4 (bases 1 to 10819)		/codon_start=1
AUTHORS	Plunkett,G. III.		/transl_table=11
TITLE	Direct Submission		/product="orf, hypothetical protein"
JOURNAL	Submitted (13-OCT-1998) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA		/protein_id="AAC77173.1"
COMMENT	On Sep 9, 1997 this sequence version replaced gi:1790660. This sequence was determined by the E. coli Genome Project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by NIH grants HG00301 and HG01428 (from The Human Genome Project and NCHGR). The entire sequence was independently determined from E. coli K12 strain MG1655. Predicted open reading frames were determined using Genemark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 (e-mail: markborov@gatech.edu). Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible (http://cgsc.biology.yale.edu). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (http://www.genetics.wisc.edu). *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.	promoter	/codon_start=1
			/transl_table=11
		promoter	/product="orf, hypothetical protein"
			/protein_id="AAC77174.1"
		promoter	/db_xref="GI:1790663"
			/translation="MAPHRRNTQOEITMKIFQRYNPLQVAKYKILFRGLYIKDVA FEFDGKILIPKVKDHLHLSVSEVNRQVMRLQTEMA"
		gene	complement(2320..3663)
			/gene="ytfL"
			/note="b4217"
		CDS	1996..2241
			/gene="ytfK"
			/function="orf; unknown"
			/note="b81; 100 pct identical amino acid sequence and equal length to YTFK_ECOLI SW: P39318"
			/codon_start=1
			/transl_table=11
			/product="orf, hypothetical protein"
			/protein_id="AAC77175.1"
			/db_xref="GI:1790664"
			/translation="MLNLSIVILCLINSAFSPMSSEISIAARIKTKLKLADENGINA QRLVMQENPGAEFTVVOIGLNAVALILGIVGDAFSPAFHSLFESRYMSAEISEQLS ILFSVLGMFTLEFADLTPKRGIMLPAEVAVALINPMKFCLYCTPLVMPFNGLANI IFRIFKLPMVRKDDITSDDIYAVAGALAGVALRQEHLEIENVELESRTVPSMTPT RENVIWFDLDEDESLAKNKVAHPHPSKFLACVMEQDIIITIGYVDSKDLNRLVANSLSA LNSGVIQNTLILVPTLTLSEALSESFKTAGDEPDYIMNEYALVGCITITLNDVMTPLMG DLVGGLEQVIARDBENSHLDGGRPIIDVYMYVDIDEPQSGNVEITIGCPMMPLMRK IPRKTDVKAFAKREYVDIDNTRIDQLLVTRIDSKATRLSPKLPDAKDREESA"
			complement(3722..3750)
FEATURES	Location/Qualifiers	promoter	
source	1..10819		
	/organism="Escherichia coli K12"		
	/strain="K12"		
	/sub_strain="MG1655"		
	/db_xref="taxon:83333"		
protein_bind	36..56		
	/note="central position to predicted promoter: -73..5"		
promoter	/bound_moiety="llyv predicted site"		
	86..114		
	/note="factor Sigma70; predicted +1 start at 4435250"		
gene	156..656		
	/gene="ytfI"		
	/note="b4215"		

```
promoter
  complement(3785..3813)
  /note="factor Sigma70; predicted +1 start at 4438844"
repeat_region
  3905..3938
  /note="factor Sigma70; predicted +1 start at 4438907"
  contains 1 REP (repetitive extragenic palindromic) element;
  complement(3986..4624)
gene
  /gene="msrA"
  /note="b4219"
  complement(3986..4624)
  /gene="msrA"
  /function="enzyme; Proteins - translation and
  modification
  /note="f212: 100 pct identical amino acid sequence and
  equal length to PMSR_ECOLI SW: P27110"
  /codon_start=1
  /transl_table=11
  /product="peptide methionine sulfoxide reductase"
  /protein_id="AAC77176.1"
  /db_xref="GI:1790665"
  /translation="MSLPDKKHLVSPDALPGHNPMPVATLHAVNGHSMTNVPDGM
  EIAIFAMGCFMGVLEFLMQLPGVYSTAGTGGTTPMTREVCSGDTGHAFAVRIYD
  PSVSYEOLLOVEHNDHPAOGROGNDHGTORSALVPLTPDDAARASLERFQAA
  MIAADDDRHITTEINATPFIYAEDDHQYLHKNPYGCIGICIGVCLPPEA"
  4662..4690
promoter
  /note="factor Sigma70; predicted +1 start at 4439826"
  complement(4687..4718)
  /note="factor Sigma70; predicted +1 start at 4439809"
  4751..4762
protein_bind
  /note="central position to predicted promoter:61"
  /bound_molety="Arac predicted site"
  4830..6563
gene
  /gene="yrfM"
  /note="b4220"
  4830..6563
CDS
  /gene="yrfM"
  /function="orf; Unknown"
  /note="o577: 100 pct identical amino acid sequence and
  equal length to YrfM_ECOLI SW: P39320"
  /codon_start=1
  /transl_table=11
  /product="orf, hypothetical protein"
  /protein_id="AAC77177.1"
  /db_xref="GI:1790666"
  /translation="WRIYRQLCVSLSCSNAVANRVLQYEGLSGLEKVRQSL
  TIESDEVPRFRARVDALREGLKALGTYPTLEFDLRPPKRGVYLAKVTPGV
  PVLIGTGVVLRGARTDKDKLDTBPATIGTVLQGDYENKFKSLTSLARKGYFD
```

Alignment Scores:

```
Pred. No.: 1.67e-17 Length: 10819
Score: 433.50 Matches: 164
Percent Similarity: 36.92% Conservative: 100
Best Local Similarity: 22.94% Mismatches: 262
Query Match: 9.17% Indels: 189
DB: 1 Gaps: 18
```

US-09-914-168-2 (1-919) x AE000493 (1-10819)

```
QY 211 AANIIeIyAlaIaIeUcIuAspIIeThrGInGSeAlaMeTasPleuAsnGlySer 230
  |||:||||| ||| |||
DB 4941 AACGTTCCGCGCCAGCTTTCTACGATTGGAAGTGAGTGACGCCAGCGCTGCTTT 5000
  |||:||||| ||| |||
QY 231 ILeProArGLeuArgGInThrAlaLeuValAlaIaIaArgAlaValGlyTyrTyr----- 248
  |||:||||| ||| |||
DB 5001 CGCGCACCGCTGATGATGCCATCCCGCAAGGTCTGAAAAGCGCTGATTAATTAACGACCG 5060
  |||:||||| ||| |||
QY 249 -----AspIIeAsPleuSerIIeIleArGAsnSerIIeGlyGluValAspAlaIleIle 266
  |||:||||| ||| |||
DB 5061 ACCATTGAATTGATCTCGCTCCACCGCAAGAAAGCGCGCAGATTAATGATCCGCAAA 5120
  |||:||||| ||| |||
QY 267 HIsAsPleuGlyGluProValTyrIIeAsPtyrArGAlaValGluValArGlyGlyGly 286
  |||:||||| ||| |||
DB 5121 GTCACGCGAGCGCTCGCTTAATTGGCGGACCGATGATGTGCGCGGCGGCGG 5180
```

```
QY 287 AlaSpAspIySAlaPheThrThrValAlaIaSpGluValProLeuLeuIIeGlyAspVal 306
  |||:||||| ||| |||
DB 5181 CGACCATTAAGACGTATTGAAATTCGTCGATCCGCCG---CCATTGGACGCGTA 5237
  |||:||||| ||| |||
QY 307 PheHIsHsGlyLysTyrGluThrThrLysLysAsnLeuIIeGluAsnAlaSerAlaGluHIs 326
  |||:||||| ||| |||
DB 5238 CTGAACGACGCGCATTTATGAAATTTCAAAAGCTTAAACACGATTTGCTCCGTAAA 5297
  |||:||||| ||| |||
QY 327 GLyTyrPheAspGlyArGtTProLeuAspArgSerValAspValIIeLeuProAspAsnThr 346
  |||:||||| ||| |||
DB 5298 GGTATTTCGATACGGAATTATTAACCAAGCGCAGCTGGCATTTGCGCTCGCCCATAAA 5357
  |||:||||| ||| |||
QY 347 AlaSpAlaSerLeuIIeTyrAspThrGlyThrChInTyrArGpHeAspGluValAlaPhe 366
  |||:||||| ||| |||
DB 5358 GCTTTCGGGATTAATGATTAATACAGTGGCAACGTTACGCTTGGCATGTGACCTTT 5417
  |||:||||| ||| |||
QY 367 PheThrIIeAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLys 386
  |||:||||| ||| |||
DB 5418 -----GAGGATCACAAAATCCG 5435
  |||:||||| ||| |||
QY 387 ArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAla 406
  |||:||||| ||| |||
DB 5436 GATGAATACCTGCAGAAATCTGGTCCGCTTAAAGAGGCGCATGATGATCAATCAAGAT 5495
  |||:||||| ||| |||
QY 407 ValArGAlaLeuSerAsnAspLeuIIeAlaThrArGtTyrPheAsnMetValAsnThrGlu 426
  |||:||||| ||| |||
DB 5496 CTGGCAACATGTAACCCGCCACCTTCTGCTACGCGCTGTTACTCGGTG----- 5546
  |||:||||| ||| |||
QY 427 IIeValPheProGluArGlyGluGlnIIeGlnAsnAspGlnValSerPheGluGlnSerSer 446
  |||:||||| ||| |||
DB 5547 GTGCTGCTCCACAA-----TTTGATTAAGCGCGC 5576
  |||:||||| ||| |||
QY 447 SerSerArGThrGluProAlaGlnValAspGluSerThrLeuGluProValIIeGluThr 466
  |||:||||| ||| |||
DB 5577 GAACAGAAATATTACCA----- 5594
  |||:||||| ||| |||
QY 467 ValGluLeuThrAspGlyIIeLeuMetAspIIeSerProIIeGluPheSerAlaSerAsn 486
  |||:||||| ||| |||
DB 5595 -----TTGCACG---GCGCG-----GTTTCGCGC----- 5615
  |||:||||| ||| |||
QY 487 LeuIIeGlnAspLysLeuAsnLeuValAlaIaIaLysAlaArGHisLeuTyrArSpmetPro 506
  |||:||||| ||| |||
DB 5615 ----- 5615
  |||:||||| ||| |||
QY 507 AspAspArgValLeuAlaIIeAsnHIsAspAspGlyValAsnArgSerIIeLeuGlyArg 526
  |||:||||| ||| |||
DB 5615 ----- 5615
  |||:||||| ||| |||
QY 527 IIeSerAspAlaValSerAlaValAlaArgAlaIIeLeuProAspGluSerGluAsnGlu 546
  |||:||||| ||| |||
DB 5616 -----CGACAGAAACACG 5630
  |||:||||| ||| |||
QY 547 ValIIeAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr 566
  |||:||||| ||| |||
DB 5631 ATC----- 5633
  |||:||||| ||| |||
QY 567 GlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArGAspGlyGln 586
  |||:||||| ||| |||
DB 5634 -----GAA 5636
  |||:||||| ||| |||
QY 587 IIeGlyLeuGlyTTrpGlySerAspThrGlyThrArgLeuValThrLysPheGlyHIsAsn 606
  |||:||||| ||| |||
DB 5637 ACCGGGCTCGGTACTCTACGACGACGCGGACCGCGGTGAAGCGACGTAAGAAAGCGCG 5696
  |||:||||| ||| |||
QY 607 LeuIIeAsnArgAspGlyTyrGlnIaGlyIaGlyLeuLeuArgLeuSerLysPlyLys 626
  |||:||||| ||| |||
DB 5697 TGGATGAACCTTATATGCTACAGCTGTACACCAACCACTACTAGTATTGTCGCGCGGCAAG 5756
  |||:||||| ||| |||
QY 627 GlyValLysLeu---TyrAlaThrLysProLeuSerHIsProLeuAsnAspGlnLeuArg 645
  |||:||||| ||| |||
DB 5757 ACCCTCGACTTCAGCATTAATAATGCGCGTGTGAGAAATCCACTGGAACAATATTATTG 5816
  |||:||||| ||| |||
```

TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS
Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A. and Blattner, F.R.	Genome sequence of enterohaemorrhagic Escherichia coli O157:H7	Nature 409 (6819), 529-533 (2001)	21074935		
2 (bases 1 to 11760)			11206551		
Perna, N.T., Plunkett, G. III, Burland, V., Mau, B., Gläsner, J.D., Rose, D.J., Mayhew, C.F., Evans, P.S., Gregor, J.J., Kirkpatrick, H.A., Postal, G., Hackett, J.F., Klink, S., Boutin, A., Shao, Y., Miller, L., Grobleck, E. J., Davis, N.W., Lim, A., Dimalanta, E., Potamovsis, K., Apodaca, J., Anantharaman, T.S., Lin, J., Yen, G., Schwartz, D.C., Welch, R.A. and Blattner, F.R.	Direct Submission				
Submitted (22-OCT-2000) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA					
Location/Organism					
1. .11760					
/organism="Escherichia coli O157:H7 EDL933"					
/strain="EDL933"					
/serotype="O157:H7"					
/db_xref="taxon:155864"					
/note="enterohemorrhagic"					
126. .866					
/gene="cysQ"					
/note="25825"					
126. .866					
/gene="cysQ"					
/function="phenotype: Central intermediary metabolism: Sulfur metabolism"					
/note="Residues 1 to 246 of 246 are 99.18 pct identical to residues 1 to 246 of 246 from Escherichia coli K-12 Strain MG1655; B4214"					
1078. .2016					
/transl_table=1					
/product="affecta pool of 3-phosphoadenosine-5'-phosphosulfate in pathway of sulfite synthesis"					
/db_xref="GI:12519213"					
/protein_id="AAG59412.1"					
/translation="MLDVCQLARNAGDAIQVYDGAKPMDVSKADNSPTADIADAAHYIMDGLRTLPDIPVISEDPRGVMORHOMVYMLVDPLDGRKEFKRNGEPTVNIALIDHGKPLRIGVVYAPYMNWYSAABEKAKEGCVKQIIVRPARPLVYISRHADALFKYEQQLGEMOTSTIGSKRCLVABEGAOALYPRRGPINIMDTAAGAAVAAACAGHVVHDMQKPLDYIPRESFLNPGFRVSY"					
1078. .2016					
/gene="25826"					
1078. .2016					
/gene="25826"					
/function="orf; Unknown function"					
/note="Residues 1 to 133 of 312 are 98.49 pct identical to residues 1 to 133 of 166 from Escherichia coli K-12 Strain MG1655; B4215"					
1078. .2016					
/transl_table=1					
/protein_id="AAG59413.1"					
/db_xref="GI:12519214"					
/translation="MDPRIRNNFICGAVELPVKSSYKTKHSNNFNNIHAFFKKDKWITSNDLSLTLRFPCATQIDKSSYQVSSSKNETYANADKFLISPLKDNQGEYMTLTKGSGEYVEIDITINIEYSPFMDKRNREPLEDNRRLYAMDILQKOSDIPKIKIYIAGISGVYALDLDDIVSKLRKRGPELOGPGGICQDCGWYIEKTIIDNPATIVESIGDIAFKTIYGGDERFRNGSSVLEDEDYSILYINLVNAGSQOVLAAQVDAIVSANLAADSDITRKERICEKRYKSTIEAFKKQQLALPLVLRKKNSE"					
complement(2079. .2633)					
/gene="ytfJ"					
/note="25827"					
complement(2079. .2633)					
/gene="ytfJ"					
/function="orf; Unknown function"					
/note="Residues 1 to 184 of 184 are 98.91 pct identical to residues 1 to 184 of 184 from Escherichia coli K-12 Strain MG1655; B4216"					
1078. .2016					
/codon_start=1					

Oy	327	GlyThrPheAspCylarGrrrPrpleuAspArgSerValAspAlIleLeuProAspAsnThr	346
Dd	6238	GGrTATTTGGATAGCGCAATTTCACAAAGCCGCATGGCGCTTGCGGTGCATTAATAA	6297
Oy	347	AlaAspValSerLeuIleTyraSprThrgLythrGlInTyraArgPheAspGluValAlaPhe	366
		::: ::: :::	
Dd	6298	GCCTTCTGGCATATATGATTATACAGAGGGCAGCTTACCAGCTTGCGCATGTGCACCTTT	6357
Oy	367	PheThrIleAspProLysThrAsnGlnLeuThrThraSprProAspLysLeuProValLys	386
		::: ::: :::	
Dd	6358	-----GAAGATCAACAATCCGC	6375
Oy	387	ArgGluLeuLeuglInLeuLeuThrValAsnMetGlyGluAlaTyraAsnLeuGlnAla	406
		::: ::: ::: :::	
Dd	6376	GATGAAATACTCGCAAAATCGTGGGCCGTTTAAAGAGGCGCATGATCGAATCGAAAAGAT	6435
Oy	407	ValArgAlaLeuSerAsnSpleuIleAlaThraArgTyraPheAsnMetValAsnThrGlu	426
		: ::: :::	
Dd	6436	CTGGCGAAGATCGAACCGCCGACTTCTGTCACCGCGCTGTAACTCGCGTG-----	6486
Oy	427	IleValPheProGluArgGluGlnIleGlnAspGlnValSerPheGluGlnSerSer	446
		::: ::: ::: ::: ::: ::: ::: ::: ::: :::	
Dd	6487	GTGGTGCTCCACAA-----TTTGATAAAGCCCGC	6516
Oy	447	SerSerArgThrGluProIacInValAspGluSerThrLeuGluProValIleGluThr	466
		:::::	
Dd	6517	GAACAGAAATATTACCA-----	6534
Oy	467	ValGluLeuThrAspLysIleLeuMetAspIleSerProIeGluPheSerAlaSerAsn	486
		::: ::: ::: ::: ::: ::: ::: ::: :::	
Dd	6535	-----TTTAGC-----GGTTCGCGC-----	6555
Oy	487	LeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyraSprMetPro	506

Dd	6555	-----	6555
Oy	507	AspAspArgValLeuAlaIleAsnHisAspAspLysAlaAsnArgSerIleLeuGlyArg	526

Dd	6555	-----	6555
Oy	527	IleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu	546
		::::: ::: ::: ::: ::: ::: ::: ::: :::	
Dd	6556	-----CGAACGAAAAACACT	6570
Oy	547	ValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyra	566
		:::	
Dd	6571	ATC-----	6573
Oy	567	GlnSerLysLysValProLeuTyraValPheValAlaSerSprLysProAlaGaspGln	586

Dd	6574	-----GAA	6576
Oy	587	IleGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAsn	606
		::: ::: ::: :::	
Dd	6577	ACCGGGGTGGTACTACGTACGACGACGTGGACCCGCGCTGAAGAGCAGCTGGAAAAAACCG	6636
Oy	607	LeuIleAsnArgSpolyTyrrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLys	626
		::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::	
Dd	6637	TGGATGAATCATACATACGGTCCACACTGCACCCAGTACCGATATTCTCGCGCCGCAACAG	6656
Oy	627	GlyValLysLeu---TyralaThrLysProLeuSerHisProLeuAsnAspGlnLeuArg	645
		:::	
Dd	6697	ACCCTGCAGCTTCAGCTATAAATGCCCGCTCTGACAGATCACCTGGAACAATATTATTG	6736
Oy	646	AlaThrLeuGlyTyrrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSer	665
		::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::	
Dd	6757	GTGCAAGGCGGTTTAAAGCCACT-----GACCTGAACGATACAGACTCTGAC	6804
Oy	666	ThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGlyTyrrParArg	685
		::: ::: ::: ::: ::: ::: ::: ::: ::: ::: :::	
Dd	6805	TCCACTACGCTG-----GTGGCTTCTGCTACTGGGATCTTCCACGCGTGGCGACGCT	6858

OY	666	ThyrisSerLeuThrArgLeuLysPheLysLeuTstHnAlaIleProGluHnrTrp	705
Db	6859	GCCATTAACTCGCTGGAGCTGTGCAC-----	6885
OY	706	GlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeu-----	723
Db	6886	-----CACTTTACTCAGGGTGTAATATTACAACACCACGATGCTTTTAT	6930
OY	724	AlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnPrometArgGlyLyr	743
Db	6931	CCTGGGGGATGATTAGCCGCACGGGTCCTCTGTGGCCCTCATGCCAACCACTGGGGCGAC	6990
OY	744	ArgGlnArgTyrSerLeuGlnValAlcLysSerSergLyLeuValSerAspAlasnmelA	763
Db	6991	TGCGAACGGTACTCTATGTGACTACTCCACACACGGCCTGGGGTTCAGATGTGATTTCCTC	7050
OY	764	IleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlatrycLysSerAsn	783
Db	7051	GTTTTCCAGGCGCAACAGCTGTGATCCGCACACGTACGATCGC-----	7095
OY	784	ArgAlaHisGlnMetThrGlyIylIleGlnAlaGlyTyrIleTrpSerAspAsn	803
Db	7096	-----CAWCGCTTTGTGACACGGCGCACCGCTGGCTGATGAACCGGTGATTTCGAC	7149
OY	804	HisValProTyrArgLeuArgPhePheHisGlyLysAspGlnSerIleArgLyltyrAla	823
Db	7150	AAAGTACCGCCCGCATCTCGTTCTTCCCGGGGCGATCGCATTCGCGGCTTACAA	7209
OY	824	HisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAla	843
Db	7210	TACAAATCTATTCGCCGCCCAAATACGCTTAACGCTGACCTGAAGAAGGGCCTCGAAGTTGATA	7269
OY	844	ValGlyThrAlaGlyTyrAsnTyrgIudPheMetLysAspLeuArgLeuAlaValDhegly	863
Db	7270	ACCGGATCGCTGGATTCAGTACACACGACGACCGAAATCGTGGGGCGGTCTTTGTTC	7329
OY	864	AspIleGlyAsnAlaItyrAspLysGlyPheThrArnsAspThrLysIleGlyAlaGlyAl	883
Db	7330	GATTAGTGGGAAGCGGTAAAGCATTTTCGCCCGACGACGACTTTAAACCGCATACCGGGGTC	7389
OY	884	GlyValArgTyrPalasSerProValGlyGlnValArgValAspValAlaThrGlyVal---	902
Db	7390	GGCGTAGCGTGGGATTCGCGGGTGGGCCAATCAACATCGATTTCGCGTACCGGGTGGG	7449
OY	903	---LysGluGluGlyAsnProIleLysIleHisPheIleGly	916
Db	7450	GATAAAGACGAACACGGG-----TTACAGTTTACATCGGT	7485
RESULT 17			
ECOW93			
LOCUS	ECOW93	338534 bp	DNA linear BCT 30-JAN-2001
DEFINITION	Escherichia coli K-12 chromosomal region from 92.8 to 90.1 minutes.		
ACCESSION	U14003		
VERSION	U14003.1	GI:1263172	
KEYWORDS			
SOURCE			
ORGANISM	Escherichia coli. Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.		
REFERENCE			
AUTHORS	Burkard,V., Plunkett,G. III, Sofia,H.J., Daniels,D.L. and Blattner,F.R.		
TITLE	Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92.8 through 100 minutes		
JOURNAL	Nucleic Acids Res.	23 (12),	2105-2119 (1995)
MEDLINE	95334362		
PUBMED	7610040		
REFERENCE	2 (bases 1 to 338534)		
AUTHORS	Plunkett,G.I.I.I.		
TITLE	Direct Submission		
JOURNAL	Submitted (22-AUG-1994) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA. Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax:		

COMMENT

608-263-7459
On Apr 17, 1996 this sequence version replaced g1:536929.
This sequence was determined as part of the E. coli Genome Project
(Frederick R. Blattner, director) at the University of
Wisconsin-Madison. Supported by award HG00301 from the NIH Human
Genome Project. The entire sequence was independently determined
from E. coli M61653; overlaps with other sequence determinations
are annotated. The start of this entry overlaps the end of the
entry ECOUW89 (U00006) by 1885 bp.

FEATURES

source

1. 338534
/organism="Escherichia coli"
/strain="K-12"
/sub_strain="MG1655"
/db_xref="taxon:562"
/map="92.8 to 100 minutes"

misc-feature

/note="This sequence comprises the following lambda
clones: DD928(EC27-278), DD930(EC30-188),
DD933(EC27-1320), DD935(EC30MM32), DD937(EC27-1070),
DD941(EC17-142), DD945(EC17-8), DD947(EC24A-34),
DD949(EC22-169), DD952(EC17-297), DD953(EC24-409),
DD956(EC17-103), DD958(EC30MM1), DD960(EC30MM2),
DD962(EC21-104), DD965(EC23A-40), DD968(EC30K60A-4pp),
DD970(EC19-202), DD974(EC19-61), DD975(EC18-233),
DD977(EC30MM7), DD980(EC27-1151), DD981(EC18-3),
DD984(EC17-101), DD987(EC18-115), DD990(EC17-136),
DD992(EC18-282), DD995(EC17-115), DD997(EC18-113); M13mp19
or Janus vectors were used for subcloning"
1. 1885

CDS

/note="1885 bp overlap with end of Genbank Accession
Number U00006 (ECOWU89)"
complement(<1..398)
/note="ORF_1326 of Genbank Accession Number U00006
(ECOWU89)"
/codon_start=1
/transl_table=11
/label_ORF_f132p
/protein_id="AAA96985.1"
/db_xref="GI:1263173"

misc-structure

/translation="MGFTTRVSEASEKKPFNFALFMDKYGFETLAIIVAIIGSLSP
EYFLTNITQIFVQSSVYLIVGGEFFAILVAGIDLSVGAIIALSGMTAKMLAGV
DPFLAAMIGVGLGALGALNGLVNMVGL"
277..397

CDS

/note="predicted bend of 75.10 degrees"
complement(377..1909)
/note="ORF_1510"
/codon_start=1
/transl_table=11
/label_ORF_f510
/protein_id="AAA96986.1"
/db_xref="GI:536931"

/translation="MATPYISMAGIGKSGEPYHAKSVNLVYVPEIHALLGENGAK
SLMKVLSGHEHPKGTITINISYNKLDHKLAAQLGIGIYQELSVDELVLLENLY
IGRLTKKICGVNIIDREMRVRAAMMLRQLKYLDEKVANLSISHOMLEIKTL
MDARVLIIMDEPTSSLTNKEVDYFLINMQLKRGTAIYISHLAEIRICDRTATV
KDGSSGCGIYSDVANDDIIVLRVMERELONRENAKENVNLAHEFVEFVTRSDR
KKVRISFSCVCGELIGFAGLVSGRTELNMLRGVDKRAGETIRLNCGDIDSPRDL
AVKGMATITERRRNGFFPNPSTIAONNAISLSLDDGKKGAMGLHEHDEQRTAENO
RELLAKCHSVNQTITELSGGNOQAVLISKVCCPEVLIIEPFRGIDVGAKEIYK
VWRLADGKXVILMVSSLPETITVCDRIAVFCBGRLOITLRDMSSEETMAVALP
08"

misc-structure

809..931

misc-structure

/note="predicted bend of 74.75 degrees"
1768..1889

promoter

/note="predicted bend of 75.87 degrees"

repeat_region

/note="promoter matrix score of 55; putative; resembles
filh-regulated promoter"
complement(1985..2021)
/standard_name="REP; repetitive extragenic palindromic
element"

CDS

/note="contains 1 REP sequence"
complement(2036..2971)

/note="ORF_f311"
/codon_start=1
/transl_table=11
/label_ORF_f311
/protein_id="AAA96987.1"
/db_xref="GI:536932"

misc-structure

3010..3121

/note="predicted bend of 73.22 degrees"
complement(3030..3953)
/note="ORF_f307"

CDS

/codon_start=1
/transl_table=11
/label_ORF_f307
/protein_id="AAA96988.1"
/db_xref="GI:536933"

promoter

/translation="MASSIOGEMNWSQSPFDSALPNIGLGLATYLRMKOGCHTENSR
IYVWLKPGULSCAPAIKDYAEHLAVSEATIVYVKLGFSGFRNLRSALDEYFSOSE
QVLPSELAEADAPQDVVKNYENTTLRTIMEGQSIIVANDEIHRARFEYQARODLYGA
GGSNAICADQVQHEFLRIGVRCQAVPDPAHIMMSASLQEDGVLLVYTHSRTSDVAA
VELAKRNGAKIICITHSYHSPILAKYADIISCSAPETPLGRNARSARILOTLDLDAF
VSAQNLNEBANINMOKTAIVDFSPGALK"
complement(3952..3979)

misc-structure

/note="promoter matrix score of 61; putative"
4151..4251

gene

/note="predicted bend of 72.06 degrees"
4279..4728
/gene="yjca"

CDS

4279..4728
/gene="yjca"
/codon_start=1
/transl_table=11
/label_ORF_o149
/protein_id="AAA96989.1"
/db_xref="GI:536934"

misc-feature

/translation="MKKIAFGCHVGFILKHEIIVAHVGERGVEIDKGMSSBRTDP
HYASOVALVAGGEVDGILICTGCGISIAKRAKGRVAVGSEPSYSAOLSRQNDPT
NVIAFSRVYGEELAKMIDAMIGAGYEGGRHQRVEALITAEORNR"
complement(4520..16191)

gene

/note="corresponds to Genbank Accession Number D90227
(ECOPHN) 1..11672; the related sequence Genbank Accession
Number J05260 (ECOPHNAO) is from E. coli B, not K-12"

CDS

complement(4739..5146)
/gene="phnQ"
complement(4739..5146)

/gene="phnQ"
/note="annotated in Genbank Accession Number J05260
(ECOPHNAO), but probably not a gene"

/codon_start=1
/transl_table=11
/protein_id="AAA96990.1"
/db_xref="GI:536935"
/translation="MPAPYPTSSPILFIHSPFLIOLPPLLKALAIYOLPAIIFKL
AAAVLISAPIVAHFHLPVAVARIYFVSPVRIYVINRLNIANGVONIGIGRNV
SVSRVAKRMSGRFISQPKETSLIITHPC"

gene

complement(5272..6030)

CDS

/gene="phnP"
complement(5272..6030)
/gene="phnP"

/codon_start=1
/transl_table=11
/protein_id="AAA96991.1"
/db_xref="GI:536936"

/translation="MSLTLLTGGAQGVPMWGECACARARSPQYRPOPCGV
KNDATITLDAGLHDLADRWSPGQOFLITVYHMDHVGGLPRLRWGADPILPVGP
DEQGDLDPEKHPGLDLSHTVEPVPVFDLQGVPLPLNHSKLTGEGYLETASHSVA
WISDAGLPEKTLKFLRNQPOVMVMDCSHPRAARAQRHCDLNTVLAALNOYIRSRV
ILTHISHQFDAMLENALPSGFEEVGDGEIYGA"


```

QY 724 ALAclValAlaValAlaHisThrValAlaAspAsnLeuValAsnProMetArgGlyTyr 743
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 134374 CTTGGGGATGATGATTAGCCGACGCGTTCTCGTGGGCTGATGCCAACCTGGGCGAC 134433
QY 744 ARGcInaArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 134434 TCCCAACCTACTCTCATCTACACTCCAAACAGCGGCTGGGCTGAGATGTCATTTCTCC 134493
QY 764 lLeAlaArgAlaGlyLleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsn 783
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 134494 GTTTCACAGCCGACAGACTCTGGATCCGACACTGTACGATGCC----- 134538
QY 784 ARGAlaHisGlyMetThrGlyGlyLleGlnAlaGlyTyrLleTyrSerAspAsnPheAsn 803
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 134539 -----CATCGTTTGTATACAGCGGACGCGTGGCTGATTGAACCGGATTTTCGAC 134592
QY 804 HlSValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerLleArgGlyTyrAla 823
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 134593 AAGTACGCGCGGATCTCGTTCTTCGCGGGGCGAGCGGACGAGATTCGTGGCTACAA 134652
QY 824 HlSAspSerLeuSerProLleSerAspLysGlyTyrLleuThrGlyGlyGlnValLeuAla 843
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 134653 TACAATATCTATCGCTCCGAATATACCCAAACGCTGACCTGAAGGGCGCTCGAAGTTGATA 134712
QY 844 ValGlyThrAlaGlyTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGly 863
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 134713 ACCGATCTCGGATACAGTACAGTCAACGCGAATAATGCTGGCGCGCTGTTTGTGC 134772
QY 864 AsPIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysLleGlyAlaGlyVal 883
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 134773 GATAGTGGCGCAAGCGGTAAGCATATTGCGCGACGACTTTAAACCGGTACGGGGTTC 134832
QY 884 GLyAlaTArgTyrPalaSerProValGlyGlnValArgValAspValAlaThrGlyAla--- 902
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 134833 GCGGTGCGCTGGGATCGCGGTGGGCAATCAAACTGATTTTGCCGTACCGGTGCGG 134892
QY 903 ---LysGluGluGlyAsnProLleLysLeuHisPheMetLleGly 916
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 134893 CATAAAGCAACACGCG-----TTACAGTTTACATCGGT 134928

RESULT 18
AX033470 AX033470 1731 bp DNA linear PAT 21-SEP-2000
LOCUS DEFINITION Sequence 3 from Patent W00047737.
ACCESSION AX033470
VERSION AX033470.1 GI:10280230
KEYWORDS
SOURCE
ORGANISM
Haemophilus influenzae.
Haemophilus influenzae
Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Haemophilus.
REFERENCE
1 (bases 1 to 1731)
AUTHORS
Ruelle,J.L. and Thonnard,J.
TITLE
Haemophilus influenzae rd outer membrane sequences used as vaccine
JOURNAL
Patent: WO 0047737-A 3 17-AUG-2000.
RUELLE JEAN LOUIS (BE) ; SMITHKLINE BEECHAM BIOLOG (BE) ; THONNARD
JOELLE (BE)
FEATURES
source
1. 1731
Location/Qualifiers
/organism="Haemophilus influenzae"
/db_xref="taxon:727"
BASE COUNT 509 a 322 c 386 g 514 t
ORIGIN

```

```

QY 243 ARGAlaValAlaGlyTyrTyrAspLleAspLeuSerLleLeuArgAsnSer---LleGlyGlu 261
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 211 CGTGTGTTGGTTATATGATCTCCGTCTTTGAAGAAACGCTCAAGCGAAA 270
QY 262 ValAspValLleLleHisAspLeu-----GlyGluProValTyrLleAspTyrArgAla 279
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 271 CCGGATTTATGATCTCATGTTACACCGGCGAGCCAAACAAATTCGGGGAGAT 330
QY 280 ValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluVal 299
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 331 GTCCAAATTGAGGGGAGAGCGCAAGATGAATAATTTGATCGCTAATAAACTTG 390
QY 300 ProLeuLeuLleGlyAspValPheHisGlyLysGlyTyrGluThrLysAsnLeuLle 319
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 391 CCA---AAGAGAGCGCTTTGGTTGTAACACCAACTTACATATTCACAAACGCGATT 447
QY 320 GluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrLeuAspArgSerValAsp 339
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 448 TCACGCTTGGCATTAATCGTGGTATTTGATGGAACTTTAAATTCACGTTAGAA 507
QY 340 ValLleLeuProAspAsnThrAlaAspValSerLeuLleTyrAspThrGlyThrGlnTyr 359
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 508 ATCAGCCCTGAACCCCATCAAGCATGTGGCGAATGTATATGATGCTGCTTAT 567
QY 360 ArgPheAspGluValAlaPhePheThrLleAspProLysThrAsnGlnLeuThrThrAsp 379
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 568 CATTAATGCAATATTAATCTTT----- 588
QY 380 ProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGly 399
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 589 ---AGCCATTCAACAAATCCGTGACGATTAATCTCAATAAATATCTTAACATCTGCG 645
QY 400 GluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAspAspLeuLleAlaThrArgTyr 419
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 646 GATCCATATTATGATATTAATTTGTGGAATTAACACGAGATTTTCATCTTAATTTGG 705
QY 420 PheAsnMetValAsnThrGluLleValPheProGluArgGluGlnLleGlnAsnAspGln 439
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 706 TTAGCTCAGTA----- 717
QY 440 ValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThr 459
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 717 ----- 717
QY 460 LeuGluProValLleGluThrValGluLeuThrAspGlyLleLeuMetAspLleSerPro 479
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 717 ----- 717
QY 480 lLleGluPheSerAlaSerAsnLeuLleGlnAspLysLeuAsnLeuValAlaAlaLysAla 499
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 718 -----TTAGTTCAGCCCTAATGTTAAT----- 738
QY 500 ArgHisLleuTyrAspMetProAspAspArgValLeuAlaLleAsnHisAspAspGlyVal 519
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 738 ----- 738
QY 520 AsnArgSerLleLeuGlyArgLleSerAspAlaValSerAlaValAlaArgAlaLleLeu 539
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 738 ----- 738
QY 540 ProAspGluSerGluAsnGluValLleAspLeuProGluArgThrAlaLeuAlaAsnArg 559
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 738 ----- 738
QY 560 LysThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSer 579
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 739 -----CATAAAGCAAAACCTGTGATGTGAGATGTATCTTTATCCA 780
QY 580 AspLysProArgAspGlyGlnLleGlyLeuGlyTyrPglYSerAspPheGlyThrArgLeu 599
    ||||| : : : : : ||||| : : : : : ||||| : : : : :
Db 781 CGTAAATAAATAATGCATGAGACTCGGTGTGGCTTGTACTGATGCGCGCTTACGGA 840

```

Oy 600 ValThrLysPheGluHisAsnLeuIleAsnArgSerGlyTyrGlnAlaGlyValAlaGluLeu 619
 Db 841 CAAATTTGGCTGGACAAACCTTTGGATTATAGCCGTGGACATAGTTTCCGTTCAAAATTT 900
 Oy 620 ArgLeuSerGluAspLysGlyValLys---LeuTyrAlaThrLysProLeuSerLys 638
 Db 901 TATCTCTTGCACCAAAACCAACTCTAAGGCAACTTATCCAAATGCCACTGTTAAAAAT 960
 Oy 639 ProLeuAsnAspGlnLeuArgAlaThrPheGluTyrGlnGlnGlnValPheGlyHisSer 658
 Db 961 CCATTAAATTTATTACTATGATTTTCCCTCGGTGGAA-----GGGGAANA 1008
 Oy 659 ThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIle 678
 Db 1009 GAGAAC-----GATACCAATACGAGAGCTGTACGTGTGCAGCGTTACGTTATTGGAA 1062
 Oy 679 GlnAsnGlyTyrPasnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThr 698
 Db 1063 AATGCCGATGGTTGGCAATATTTTGGCGGACTTGTACCGCATACGACAGATT---ACA 1119
 Oy 699 GlnAlaProGluTyrThrPoloInaspLeuProValAspPheValAsnGlyLysProSer 718
 Db 1120 CAACGGGAT----- 1128
 Oy 719 GlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsn 738
 Db 1129 -----ATCAGTCGATAAACCTTA-----CTTCTTTAT 1155
 Oy 739 PrometArgGlyTyrArg----- 744
 Db 1156 CCAACTGTGGATTACTTCGCACTCGATTTAGCTGGTGGTCCCTTTGCCACTTTGGGGCAT 1215
 Oy 745 ---GlnArgTyrSerLeuGlnValAlaGlySerSerGlyLeuLysSerAspAlaAsnMetAla 763
 Db 1216 GTGCAGAAAAATTTACTTTGATTATTAAGCAACAGAAATTTGGCATCGAATCTCTTTATA 1275
 Oy 764 IleAlaArgAlaGlyIleSerGlyValLysSerPheGlyAspAsnAlaTyrGlySerAsn 783
 Db 1276 AAAGTCAGAGCATTCAGCGCGGTGGTGTATGCAAGAAAT----- 1320
 Oy 784 ArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsn 803
 Db 1321 -----CATCGATCGTTGCTGCTGCTCAAAATCGGGATTTTACATACAAAGATTTGAA 1374
 Oy 804 HisValProTyrArgLeuArgPhePheAlaGlyIleAspGlnSerIleArgGlyTyrAla 823
 Db 1375 AAAATTCGGCTACACTCGCTTCTTCTGCTGTCGATCGATCGATCGCGGGTTACGCG 1434
 Oy 824 HisAspSerLeuSerProIleSerAspLysGlyTyrIleThrGlyGlyGlnValLeuAla 843
 Db 1435 TATATAAAAAATTTGCCCTTAATAATGAAATGAAATATGGTTGGCGCTCACGCTTCTT 1494
 Oy 844 ValGlyThrAlaGluTyrAsnTyrGluThrPheLeuLysAspLeuArgLeuAlaValPheGly 863
 Db 1495 ACCACTCTTTAGAAATCATATCATACGTTTATCCAAATTTGGTGGCGCACTTTTGCA 1554
 Oy 864 AspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyAla 883
 Db 1555 GATATGTGATTAGCTGCCGATATATTAACACAGCAAAAGACGTGTTATGACAGAGCTT 1614
 Oy 884 GlyValArgTyrPalaSerProValGlyGlnValArgValAspValAlaThrGlyValLys 903
 Db 1615 GGTGTGCGTGGCGCATCCCACTGAGTGGTGCATTAATTTGATATGACACACCCATTTGCT 1674
 Oy 904 GlnGluGlyAsnProIleLysLeuHisPhePheIleGly 916
 Db 1675 GATTAAGATTAACAGCAAAAAATTTTCAATTTTATCATCGGA 1713

RESULT 19
 LOCUS AE012531/c
 DEFINITION Xanthomonas campestris pv. campestris str. ATCC 33913, section 439
 of 460 of the complete genome.

ACCESSION	AE012531 AE008922
VERSION	AE012531.1 GI:21115353
KEYWORDS	
SOURCE	Xanthomonas campestris pv. campestris str. ATCC 33913.
ORGANISM	Xanthomonas campestris pv. campestris str. ATCC 33913 Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group; Xanthomonas.
REFERENCE	1 (bases 1 to 11870)
AUTHORS	da Silva,A.C.R., Ferro,J.A., Relnach,F.C., Farah,C.S., Furlan,L.R., Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C., Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J., Chambergio,F., Ciapina,L.P., Ciccarelli,R.M.B., Coutinho,L.L., Cursino-Santos,J.R., El-Dorriy,H., Faria,J.B., Ferreira,A.J.S., Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C., Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite Jr.,R.P., Lemos,E.G.M., Katsuyama,A.M., Kishi,L.T., Leite Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Melandris,J., Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M., Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A., Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spindola,L.A.F., Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and Kitajima,J.P.
TITLE	Comparison of the genomes of two Xanthomonas pathogens with differing host specificities
JOURNAL	Nature 417 (6887), 459-463 (2002)
MEDLINE	22022145
PUBMED	12024217
REFERENCE	2 (bases 1 to 11870)
AUTHORS	da Silva,A.C.R., Ferro,J.A., Relnach,F.C., Farah,C.S., Furlan,L.R., Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida Jr.,N.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C., Camargo,L.E.A., Camarotte,G., Cannavan,F., Cardozo,J., Chambergio,F., Ciapina,L.P., Ciccarelli,R.M.B., Coutinho,L.L., Cursino-Santos,J.R., El-Dorriy,H., Faria,J.B., Ferreira,A.J.S., Ferreira,R.C.C., Ferro,M.I.T., Formighieri,E.F., Franco,M.C., Greggio,C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locati,E.C., Machado,M.A., Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Melandris,J., Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M., Okura,V.K., Oliveira,M.C., Oliveira,V.R., Pereira Jr.,H.A., Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spindola,L.A.F., Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos Santos,M., Truffi,D., Tsai,S.M., White,F.F., Setubal,J.C. and Kitajima,J.P.
TITLE	Direct Submission
JOURNAL	Submitted (28-NOV-2001) Departamento de Bioquímica, Universidade de Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900, Brazil
FEATURES	Location/Qualifiers
source	1..11870
	/organism="Xanthomonas campestris pv. campestris str. ATCC 33913"
	/strain="ATCC 33913"
	/db_xref="ATCC:33913"
	/db_xref="taxon:190485"
	/note="pathovar: campestris"
gene	480..2099
	/gene="XCC4080"
CDS	480..2099
	/gene="XCC4080"
	/note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
	/codon_start=1
	/trans_table=11
	/product="polyvinylalcohol dehydrogenase"
	/protein_id="AA043301.1"
	/db_xref="GI:21115354"
	/translation="MSSIDPFAFAASQIWRSPGNSLPFASPNWPPAGGCFYMSRFPEGKINRTAPTLKPAWTLTTTDDISATPSVEEGCTLYVPDGMGTLYSDYOSKRPNNQAKLSIDTGNANSVTRNPATPSGSLIIVGDQASCTVMAIDKTKRLMWSVAESNPQDA RITSSPIYGNRIYVIGSSGDMGMIANHFHSFGSVAIDALNKGKAWQCFYNPDEG

TCASVWGOAIDPISRLYFGTCNNYNIIPAGVNCVSNARSTRGSELTIVODELNCMA
DNVYDVALNLDITGKLMADRAQGYDAMNLSICIVAPNTGICPNTQATNLITGPDYDFG
GAGVOLFVWBRGKPOOLVGAQKSGIYAPMARTCAKVMSTQVGGTGGVEGTA
FDPYKQVYVAIINSSNATYTLGPDNTVSHNGSNMAALDAATGKIKMOVKVGVNSV
PNVPAGGGSLTIVSPNLLYAGSMACNMVALDITGATJLMNVAITGSVSISSPAVVGAL
YWGAGCFRPNCTATSNQIKFVPSGAAK"
/complement(2282..2587)
/gene="XCC4081"
/complement(2282..2587)
/note="XCC4081"
/note="putative: ORF located using Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA043302.1"
/db_xref="GI:21115355"
/translation="MAIARVVICLGRSRMNIMAAFRPAMTAPEROLLHYFADYE
GYCDRPLMGOLAVSEVPADGVDAELVCLARRRAEAHGLEAGQIRLCQVRL"
/complement(2490..2930)
/gene="arod"
/note="XCC4082"
/complement(2490..2930)
/gene="arod"
/note="identified by sequence similarity: putative: ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="3-dehydroquinate dehydratase"
/protein_id="AA043303.1"
/db_xref="GI:21115356"
/translation="MILRGPEASSLVGRPOLPAVITTOLYRCARDAGKYALRACG
SEQLDLRVAGAKMETALIDPSCVDVSRVHLVGLAIPVEAHDDSTDRPERC
LPAGIGQVVAIVRGCAQSYLGLLEIALEHLCGTBIQGDVHYGT"
/complement(2950..3372)
/gene="XCC4083"
/complement(2950..3372)
/gene="XCC4083"
/note="XCC4083"
/note="putative: ORF located using Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA043304.1"
/db_xref="GI:21115357"
/translation="MCILILRGPSDAIPVAPLPCACGRDLTLCADLASIADLQA
APLAVITPHDAIPRAYAMSLGIAACPLALHLMRVAA"
/complement(3616..7572)
/gene="XCC4084"
/complement(3616..7572)
/gene="XCC4084"
/note="identified by sequence similarity: putative: ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA043305.1"
/db_xref="GI:21115358"
/translation="MALAGASAGCYCARLERPRFPAVALHRRHGEVNTTPTSPA
PRPFYRRRRFVSGGLVILGLIALALVLAQVTLVPLIIDAEGRTPOLR
YGSVEVPAGPLTLANVDFRQDIFHTAERYLLEPDLRPLGRKLOLAVOVSNATN
LGKSDPEPTLPSPESILPQINVLAIADHIGVINDLRITLOOQPIITLHROGGLTVA
TGELTRNLVYTDKGRPLRNDVYPSSEYRADLITRATVLAARGRTPASLSGLVARGD
LDKMEVALIAGRAPALRASLVFTGRDPTMAFNATVLTSLIPAPAGETPPATP
IALDLQASGKADLRGVSVOGELSATLQPSHSLQOVLTVEPLIIDAEGRTPOLR
GSADEFDADNASFRFSVNASGLRTPADPATPDAPAVPELTDRILGLAKMAA
IGRATVORGOOAEIVPDSRGNDORAKOYAKTPGSLDTGCVNAPALOMDLTA
OLAKFPCYFVGMWGNLSGOVASKRQLPAPAGVSGCYEATADIPITLQGLRATL
SANKFALRGTQEGQDLALGSSRTDKKGVGODLDAOLQPLQDLDLLEGATGV
RGVLQVNRDAPVITADLACTGLKMNYSQASVSLRRLPWSGSDQLALQGTIAEA
RGVLDSVVRVAGVLEALRLDIANMSVSAQSVLDRRLRNGRANGOGVSTLRIPAKG
DAMALROPACSTEGAAFTLSTGLAATGALCALAMPEGMNVHGCALPLSLVOP
WLPROGROIYLRGLSLDSGFKPRGNAMESPRASSEGCVRLGEGENYVAACVAPNR
GELLRYDOFSVOADFTTPOOTGKLCVGTGQNGFADAKINTGMDAYAPLNGELYLMKSR

LYWLELVADVVRPKGLVEGHVSLRGTRODKPLLGDATLSDEFATYPSMGLTLLSEKCG
REDALPDGSAKITASAKSGDGLINIDGISTWGTSTPLLINTIRGENVLAINTSELITV
ANDPQFGIVDTTMOLEKGVTPVPEADVIDERDRGTSVEDVYVDPVPEAPASPL
DMDLIVLGDCKMNSGEFGKGLSGOMQIRARPRGMTANGDIDVGRYKAYGODLITI
SRGQLTMMNNIYSDPRVSRARERKTDYTAGIDVSRASPRADVWSEPMASQSESV
YLVGLSTASSDETOQVSAASAALSAGSSLIASOTIGKGLIDEGAVSESSVGSV
GFGKLSPLKLYGVGVSMVSGSVTLTKYLSRGPVFEASSTVETKGSIMNREK"
/complement(7470..9410)
/gene="XCC4085"
/complement(7470..9410)
/gene="XCC4085"
/note="XCC4085"
/note="identified by sequence similarity: putative: ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AA043306.1"
/db_xref="GI:21115359"
/translation="MPHVRPRDLPCEDLCMYVHDVAYMTGHALSTPPRRPNRS
RSSRYPAPMRFLIARVAPVFLICILAGSAFAGTIDKIEKIGLDDGDDAEKMEINIS
LSLYEAVKEGCEESRLEYLLAOEROTREALPFGYSPETIDIAAPRDESLTVITV
NNGEPVVRQSHISITGMAEODRYLQDOKOPEPRGQVSHROYEASVRLTRAE
RGYPADFTQRRVAVTRAHAADIDILMDSGRYMDGYRFPDYKDFRPLFEPYVWL
BESYHEGKLDLRESLTKDYESTIDIDPKPEADDOGRVYVDKTKTRATVITSG
LSYSGSGAGVGRGVERRVNSRGMKDQDLYDAQRKSLTTSYRVPATRMLDGVTYA
SARLYDEOTERYIDLRVNKLTVGRSGQINRWASAIASINALRBRMFSGSDPTDAYE
TSTLYPOLQANVYVNDRLRFPSSGOMFIRGAGEGSDPTNFGOLGOLRMFLGA
GNGRYILREGEQTTWSDLVAMPSPSLREFAGAPNSIRRYAREVOPPRPKDEFVAG
AKNVYASAEYEHYKLGWPMGGAFAVFDSSASARDPDMWTGIGFGLRMWSPVGPVAVD
IAHGLNDPDSOFOLYIDIGANL"
/complement(9598..11694)
/gene="g1ys"
/note="XCC4086: g1ys(B)"
/complement(9598..11694)
/gene="g1ys"
/note="identified by sequence similarity: putative: ORF
located using Blastx/Glimmer/Genemark"
/codon_start=1
/transl_table=11
/product="g1ycyl-tRNA synthetase beta chain"
/protein_id="AA043307.1"
/db_xref="GI:21115360"
/translation="MSEQPLILIECTEELPYKALPGLQAQAFEDCYLACLERGAVT
KQDAKPLSTRVAVLLPGVATEQEPQRESEVGPYLVNLINDGKFTRLALPAAAG
IDMTLERTSDKGEFVIRAVTPGAQAALALPEILIREAIAMPTEKPRMGAHEYAF
ABPVOMLVLFQDVTYIPAILGVRGDRIRGRFHMDDGIIAALAGDYIDALRAHVL
VADARRARIVEENDVAAARQAGSARISQNDQOVNVNLEVMESAVLCSFEERLAPVO
EALITEMINOKFPPVLDGGRKLTBOFTGINIVSDVDEVAKGVRRVIRPPADAKF
FPDEDLKGLTAMGAGIASVYQOAKLTGADAVARAAIAEAIVQVGDPPQARAA
ELAKNDLOSRYNNEPDELQIAGRIHAKAGPSESLATIDEAYOPRRGDDITLUSPL
GKVLAIARLDTLAGFAAGLKPNGKDFALRRNALGLARTVIESGFDLPLKLDV

Alignment Scores:

Pred. No.:	8,56e-14	Length:	11870
Score:	375.50	Matches:	161
Percent Similarity:	35.298	Conservative:	86
Best Local Similarity:	23.008	Mismatches:	232
Query Match:	7.948	Indels:	201
DB:	1	Gaps:	21

US-09-914-168-2 (1-919) x AE012531 (1-11870)

QY	235	ArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp-----IleAspLeu	252
DB	9032	CGGCAAGACCGCGC---GAGGCACTGGAGCGGTTGATTATTAACCGCCACACATCGACATC	8976
QY	253	SerIleIleArgAsnSerIleGlyIleValAspValIleIleHisAspLeuGlyIlePuo	272
		:::	:::
DB	8975	GCGCGCCGCGCGGAGCGGAGCGCCCTGACCGTGATGATCAGCGTGAACCGTGGCGGCGC	8916
QY	273	ValTyrIleAspTyrArgAlaValArgGlyGluGlyIleValAspAspIleAlaPhe	292
		:::	:::
DB	8915	GTCGGGTCGCGGACGTCATATTTTCGATACCGGCTGGCGCGGACGACGACGCTATCTG	8856

OY 293 ThrThrValAlaAspGluValProLeuLeuIleGluValAspValPheHisGlyLysTyr 312
 DB 8855 GGGCAGGACTCAAGCGATTGCGAGCCCGGAGGCGAGGCTTCCAGCCATCCACAGTAC 8796
 OY 313 GluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArg 332
 DB 8795 GAAGCAGGAGGTGGCGGCTGACCCCGGCTCGGCGGAGCCGCTATTTCGATGCCGAC 8736
 OY 333 TrpLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIle 352
 DB 8735 TTCACCCAGCGCGGCGGCGGTGACCGCGTCCGAGCATGCCGCGCATGATCTGAC 8676
 OY 353 TyrAspThrGlyThrGlnTyr-----ArgPheAspGluValValPhePhe 367
 DB 8675 TGGGACAGCGCGCGCGCTTACGACATGGGACCGCTGCGCTTTCGATTCATTTTC-- 8619
 OY 368 ThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArg 387
 DB 8618 -----CGCCCG 8613
 OY 388 GluLeuLeuGluGlnLeuLeuThrValAlaAsnMetGlyGluAlaTyrAsnLeuGlnAlaVal 407
 DB 8612 GGGCTGTTCGAGCGCGTGGTATTGGAGGAGGCGAGCTACTACACGAGGAGGAGCGT 8553
 OY 408 ArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIle 427
 DB 8552 GATCGGCTGGCGGAGTACTGACGACCAACGCTGATTTACACACATCGACATCGAC-- 8496
 OY 428 ValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSer 447
 DB 8496 ----- 8496
 OY 448 SerArgThrGluProAlaGlnValAlaAspGluSerThrLeuGluProValIleGluThrVal 467
 DB 8495 -----CCCAAGCGGAGGAGGAGCGGAGGAGGCGCGGCTG-----GACCTC 8448
 OY 468 GluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeu 487
 DB 8447 AACTGACC----- 8439
 OY 488 IleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAsp 507
 DB 8439 ----- 8439
 OY 508 AspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIle 527
 DB 8439 ----- 8439
 OY 528 SerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluVal 547
 DB 8438 -----CGCGCC----- 8433
 OY 548 IleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGln 567
 DB 8432 -----AAGCCGACC-----GTGATAC 8415
 OY 568 SerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIle 587
 DB 8414 TCC----- 8412
 OY 588 GlyLeuGlyTyrPheLysSerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeu 607
 DB 8411 GGTCTGAGCTATGGCAGCAGAGCGGCGGTGTGGCGGCGGCGGCGGAGCGCGCTAT 8352
 OY 608 IleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGly 627
 DB 8351 GTCAATTCCGCGCGGCGCAAGATGACAGCAGCAGCTGACTACGAGCAAGCCGAAAGAGC 8292
 OY 628 ValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThr 647
 DB 8291 CTGACCAACAGCTACCGGCTGCGCGCTTCGCTGCTGATGGTGTGATACCGCGCTCG 8232

OY 648 LeuGlyTyrGlnGlnIleValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArg 667
 DB 8231 -----GGCGGCTGTACAGCAGCAGCAGACCGAATACATGCACCTG-----CCG 8190
 OY 668 ThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrPheAsnArgThrTyr 687
 DB 8189 AACGTCAAACTCACCGGCGCGGAGTGGCGAGATCAACAGCAGCGGTGGAGTGGCGCGC 8130
 OY 688 SerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGlnAsp 707
 DB 8129 TCGATC-----AACGCTTCGCGGAGCGGTGGCTG----- 8100
 OY 708 LeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeu----AlaGlyVal 726
 DB 8099 -----TTACGACAGTGGCAGCGCATTTCCAGGATCCGCTGTACAGACATCCACG 8052
 OY 727 AlaValHisLysThrValAlaAspAsnLeuValAsn-----Promet 740
 DB 8051 CTGATCTATCCGACCTGCGAGGCAACTACGACGTCAGATGACCGGCTGTCCCGCGC 7992
 OY 741 ArgGlyTyrArgGlnArgTyrSerLeuGluValAlaGlySerSerGlyLeuValSerAspAla 760
 DB 7991 ACGGCGTGTCCGGGCGAGTGTTCATCCGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7932
 OY 761 AsnMetAlaIleAlaArgAlaGlyLysSerGlyValTyrSerPheGlyAspAsnAlaTyr 780
 DB 7931 AATTTCGGTACCTTACCGCGCAGCTACGCGTGTCTTGGCGCGCGCGCGCGCAATTCGGCGG 7872
 OY 781 GlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
 DB 7871 GTGATCTGTGCGTGGC-----GAAGGTGGCACCCATCCGTGACACG 7833
 OY 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyLysPheGlnSerIleArg 820
 DB 7832 GATCTGTGGCGATGCGCGCGGAGCTGCGGTCTTGTGGCGGTGGCCCAACAGCATCCGC 7773
 OY 821 GlyTyrAlaHisAspSerLeuSerProIleSerAspLys-----GlyTyrLeuThrGly 838
 DB 7772 GCGTATGATTCGCGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGT 7713
 OY 839 GlyIleValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPheMetLys-----Asp 856
 DB 7712 GCCAAGAACGTGTGTCACCGCGCGCGCGCGCATACGAGCAT---TACCTCAAGCGCGGCGCG 7656
 OY 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyThrTrpAsnAsp 876
 DB 7655 TGGCGCGGCGCGGTGCTGCTGATACCGGAGCTGCTTGATGAC-----ACCCCGGAC 7602
 OY 877 ThrLysIleGlyAlaGlyValGlyValArgTyrPheAspProValGlyGlnValArgVal 896
 DB 7601 TGGCGCAGCGCATTTGGCTTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCGGTG 7542
 OY 897 AspValAlaThrGlyValLysGluGluGlyAsnProIleLysLeuHisPhePheIleGly 916
 DB 7541 GATATTGGCGCAGCGCTTGAACGACCGCGATGCGAGTTCAGCTTCATCATGACATCGCGG 7482
 RESULT 20
 AEO12071/c 10839 bp DNA linear BCT 29-MAY-2002
 LOCUS AEO12071 10839 bp DNA linear BCT 29-MAY-2002
 DEFINITION Xanthomonas axonopodis pv. citri str. 306, section 449 of 469 of
 the complete genome.
 ACCESSION AEO12071 AEO08923
 VERSION AEO12071.1 GI:21110627
 KEYWORDS
 ORGANISM Xanthomonas axonopodis pv. citri str. 306.
 Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
 Xanthomonas axonopodis pv. citri str. 306.
 REFERENCES
 1 (bases 1 to 10839)
 AUTHORS da Silva,A.C.R., Ferro,J.A., Rehnach,F.C., Farah,C.S., Furian,L.R.,
 Quaggio,R.B., Montello-Vitorello,C.B., Van Sluys,M.A., Almeida
 Jr.,N.F., Alves,L.M.C., do Amaral,M.C., Bertolini,M.C.,
 Camargo,L.E.A., Camarotte,G., Camavanan,F., Cardozo,J.,

[illegible]

LPOINPLAIQADKIDVONLRITOLQPMIVLHKVQGLDVGELTRDVIDTDMG
 DFRHGDYLDPNDRADLDTAVLPAARGPASTGLAVAGDIDKMLAVLAGRAPL
 OASLFTGRDPPWARKAVPEALDTSLIPAAGSANOAPPAITAINOASGKNAPL
 HGLTKGELSATLOPSHIALADOVLVEPIVITFERBTRTKRADRDONPSFRA
 VNASGRFTTPADPAPADPLVFEVLADGALVAGLMAALIGRAVDEQOAEV
 PDSRGNDRGAKQVQAKTPGSLDITGEVAMAPALQMDAAKADPEYFAGANG
 NLSGKIASKRQRLPAPAGVSPGEFAELPSLTGQRLVSGALNGKFAFGQGBE
 LQALNSRINAKGVDOLOLAAQLOPLDLPDLPATGVROGLDVSQRDADIT
 ADIAGNLRMDYTSQONLSRGLPWRGSDGALIGTAIEAGVLDVSVQARAGAE
 ALRDADJANSASVALQGVRRNGRQCOVATIRIAPAKGANSIROAQBSTDA
 APTLSTPCALATGALCASANPREGVAVHGLPLSLVQWPLPKOEGKOITLREGL
 SLDGSKFRPGNAMESLRIASPEGJRLGETRAAAGNPNREGLITDFOVSADPT
 QOOIOGKLGIFQAGFVDAKFNMGDAVPLNGELILNMSRLVLELVIADVVRPG
 LVEGHVSLRGTDRKPLIGDGLSDTFAEYPSMGLTSEKREGRDLPDSAKITTA
 KSGPGLTVGSGLSWFGTSTPLLTINRGLVAVNTESELINPDMQFITDNTML
 RGVVYPEADIDLERLDGTSVEDVYVLDVPEQORAPLMDLAIVGDKYMSG
 FGLKGLSGOMQIRARPREMTANGCIDVGRKATQODLITRGOLITNNVYSDER
 VSLRAERKIDVYAGIDVSGRASPRADVSEPAMSGEALSTVLGRGLSTASDET
 OOVSAASALASASSLIASQIGAKLGIDEGVSOSSLGVSVGKYLSPKLYGVYGV
 SMVGGSVLTKYLLSKRFDIEASSSTIETKGSVNMWERK"

gene
 complement (8710..10491)
 /gene="XAC4204"
 complement (8710..10491)
 /gene="XAC4204"
 CDS
 /note="identified by sequence similarity: putative; ORF
 located using Blastx/Glimmer/Genemark"
 /codon_start=1
 /transl_table=11
 /product="conserved hypothetical protein"
 /protein_id="AAM39039.1"
 /db_xref="GI:21110635"
 /translation="MLKVAFLFLCTLEVPLOAVARATIDKIEIKLDDGDAEMIE
 NIOVSLIYAVGKEGESRELEYLAOAEOTREALEPFGYSPTITLAPRAGDKYT
 VITTVRGEPVRYTHSITGMAEDORVIGODIKOEPREGGVESHPOVAXSVRT
 RRLAEGRPADPFTORVATRAEHADIDLMDSGRVMKRYRVDYFRDLEPR
 LVTWDEGSTRHEKRLRLKSLDTFTSTIDQPEEKEDQCVADVDKLTKAKRT
 VVTSGLSYSGSAGVGVGVERRVYVNAHGKMDTQDLYAONRSLSITTSYVPAKRLD
 GAVTASARLIDQTEYIDLNRNKLGTGRSQIENRWSAISINLARIRWPFSSQDRE
 GAVETSTLIVYPOLOANYVVDRLFPRGSVQOMFTRGAEAGSDTNFGOLR

Alignment Scores:

Pred. No.: 7.75e-13 Length: 10839
 Score: 359.50 Matches: 158
 Percent Similarity: 35.54% Conservative: 89
 Best Local Similarity: 22.73% Mismatches: 257
 Query Match: 7.61% Indels: 191
 DB: 1 Gaps: 22

US-09-914-168-2 (1-919) x AE012071 (1-10839)

Oy 235 ArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIle 254
 Db 10272 CGGCAGACGCGT---GAGGCCCTGGAGCCGTTGGCTATTACTGCGCCACCATTAACGCTG 10216
 Oy 255 IleArgAsnSerIleGlyGlu---ValAspValIleIleHis---AspLeuGlyGluPro 272
 Db 10215 GCCGCACACGCGCGGCGGACAAAGTGACGCTGTGATCAGCGTGATCGTGGCGACGCG 10156
 Oy 273 ValTyrIleAspTyrAlaGluValAlaArgGlyGluGlyAlaAspLysAlaIle 292
 Db 10155 GTGCGGGTGGCGACCTCGCATATCTCCATACCGGCTGGCGGACGACGCGGATCTG 10096
 Oy 293 ThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisIleGlyLysTyr 312
 Db 10095 GGGCAGATCTCAAGCAGTTTGGACCGCGGACGAGGCGGAGGCTTTCAGCCATCCGACGTAC 10036
 Oy 313 GluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisIleGlyTyrPheAspLysArg 332
 Db 10035 GAAGCCGACGAGTCCGATCACCAGGCGGTGGCGGAGCGCTGGTACTTGATGAGCCGAC 9976
 Oy 333 TyrLeuAspArgSerValAlaSerValIleLeuProAspAsnThrAlaAspValSerLeuIle 352
 Db 9975 TTCACCCAGCGCGCGTTGGATCAGCGGTGCGAGCATGTGGCGGCACATCATCTGAAT 9916

Oy 353 TyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspProLys 372
 Db 9915 TGGGACAGCGCGCGCGCTACACATGGGACAGGTGCGCTT----- 9874
 Oy 373 ThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeuGln 392
 Db 9873 -----GACTACAGCTACTTCTC-----CGGAGATGGCCTGTTTCAGCCG 9838
 Oy 393 LeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsn 412
 Db 9837 CTGGGTACTGGGAGGAGGACGCTACTACACGAGGCAACCTTGACCGCATGGCGGAG 9778
 Oy 413 AspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGlnIleValPheProGluArg 432
 Db 9777 TCGGTGACCAACGCTGGATTACTTCCAGCACATCGATATCCAA----- 9736
 Oy 433 GluGlnIleGlnAsnAspGlnValSerPheGlnGlnSerSerSerArgThrGluPro 452
 Db 9735 -----CCCAAGCC 9727
 Oy 453 AlaGlnValAspGluSerThrLeuGluProValIleGlnThrValGluLeuThrAspGly 472
 Db 9726 GAAGAGCGGAGACGACGCGCTGCTGCGGTG-----GACGTCAAACTCAC----- 9679
 Oy 473 IleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492
 Db 9679 ----- 9679
 Oy 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeuAla 512
 Db 9679 ----- 9679
 Oy 513 IleAsnHisAspAspGlyValAlaAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532
 Db 9679 ----- 9679
 Oy 533 AlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
 Db 9678 -----AAG 9670
 Oy 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValPro 572
 Db 9669 CGCAGC-----GTCTACACCTCAC----- 9652
 Oy 573 LeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrGly 592
 Db 9651 -----GCTTGAGCTACGCG 9637
 Oy 593 SerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArgAspGly 612
 Db 9636 ACGCAAGCGCGCGCGGTGCGCGGTGCGGAGCGCGGCTGACTGTACGACCGCGCGG 9577
 Oy 613 TyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeuTyrAla 632
 Db 9576 CACAAGATGACACACGACGCTGATTAACGCGAGAACCCGACAGCTGACACCACTTAC 9517
 Oy 633 ThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGln 652
 Db 9516 CGCGTGGCGCGCTTCCGTTGCTGATGGCTGACACGCGCTG-----GCG 9469
 Oy 653 GluValPheGlnHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHisGln 672
 Db 9468 CGGCTGTACGACGACGACGACGACGATGACGACCTG-----GCGACGCTCAAACTCAC 9415
 Oy 673 IleSerArgSerIleIleGlnAsnGlyLysTyrAsnArgThrTyrSerLeuArgTyrArg 692
 Db 9414 GGCACCCGACGCGCGCAGATCAACAGAGCGCTGAGCGGATCGCTCGATC----- 9364
 Oy 693 LeuAspLysLeuLysThrGlnAlaArgProGluTyrThrGlnAspLeuProValAspPhe 712
 Db 9363 -----AACGCTTGGCGCGGCTTGGCGT-----TTC 9337

```

Oy 713 ValAsnGlyLysProSerGlnGluAlaLeuLeu---AlaGlyAlaAlaValaHisLysThr 731
    ::::|
    ::::|
Db 9336 AGCAGCGGCACGATTCGAGAGCGCGGTGACAGACGCTCCACACGATCTATCCGCA 9277
    ::::|
Oy 732 ValAlaAspAsnLeuValaAn-----ProMetArgGlyTyrArgGln 745
    ::::|
Db 9276 CTGCAGCGCCACGACGATGCTGCAGCAGCGCGCTGTTTCCAGGACAGCTGCTCGCGC 9217
    ::::|
Oy 746 ArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAla 765
    ::::|
Db 9216 CACATGTTCTTCGGCGCGCGCGCGAGGAGTCCGCGCTCCATACACATTCGCGCACCTG 9157
    ::::|
Oy 766 ArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAla 785
    ::::|
Db 9156 TATGCCACGCTGCGCTGTTCTGCGCGCTGCGGACAAATGCGCGTTGATCTCGCGCGG 9097
    ::::|
Oy 786 HlaGlnMetThrGlyGlyIleGlnAlaGlyTyrIleThrPheSerAspAsnPheAsnHisVal 805
    ::::|
Db 9096 -----GAAGCGGCGACACACCTGGACGACGACGCTGCGCGATG 9058
    ::::|
Oy 806 ProTyrArgLeuArgPhePheAlaGlyAspGlnSerIleArgGlyTyrAlaHisAsp 825
    ::::|
Db 9057 CCGCCGAGCGCTGCGCTGCTGCGCGGCGGCGACACGATTCGCGGTATGCGGTCCGC 8998
    ::::|
Oy 826 SerLeuSerProIleSerAspLys-----GlyTyrLeuThrGlyGlnValLeuAla 843
    ::::|
Db 8997 GAAGTCCCGCCGACGCGCCGACGCGGACGAAATTCGCGCTGCGTGCAGAACGCTGCTC 8938
    ::::|
Oy 844 ValGlyThrAlaGluTyrAsnTyrGluPheMetLys-----AspLeuArgLeuAlaVal 861
    ::::|
Db 8937 ACCGCGACGCGCGCATATACGACGAT---TACCTCAAGGCGCGCGCGCGCGCGCGGTG 8881
    ::::|
Oy 862 PheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAla 881
    ::::|
Db 8880 TTGCTGCGATGCGCGCGCGCGCTTCGATGAC-----CGCGCGGACGCGCGCACCGGTATG 8827
    ::::|
Oy 882 GlyValGlyValAlaArgTyrPheAlaSerProValGlyGlnValArgValAspValAlaThrGly 901
    ::::|
Db 8826 GGCTTCGGCGCTGCGCGCGCGCTGCGCGGTGCGCGCGCGCGCGCTGCGATATCGCACATGGC 8767
    ::::|
Oy 902 ValLysGluGlnGlyAsnProIleLysLeuHisPhePheIleGly 916
    ::::|
Db 8766 TTGAACGACCCGACCTCCCAATTCAGCTGTACATGATATCGCG 8722
    ::::|

RESULT 21
LOCUS AY010120 20310 bp DNA linear BCT 13-JUN-2002
DEFINITION Xanthomonas oryzae pv. oryzae xanthomonadin biosynthetic gene
ACCESSION AY010120
VERSION AY010120.2 GI:15055553
SOURCE
ORGANISM Xanthomonas oryzae pv. oryzae.
            Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
            Xanthomonas oryzae pv. oryzae.
            Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
            Xanthomonas.
REFERENCE 1 (bases 1 to 20310)
AUTHORS Goel,A.K., Rajagopal,L., Nagesh,N. and Sontli,R.V.
TITLE Genetic locus encoding functions involved in biosynthesis and outer
        membrane localization of xanthomonadin in Xanthomonas oryzae pv.
        oryzae
JOURNAL J. Bacteriol. 184 (13), 3539-3548 (2002)
MEDLINE 22053219
PUBMED 12057948
REFERENCE 2 (bases 1 to 15118)
AUTHORS Goel,A.K., Rajagopal,L., Nagesh,N. and Sontli,R.V.
TITLE Direct Submission
JOURNAL Submitted (18-OCT-2000) C.C.M.B., Uppal Road, Hyderabad, A.P. 500
        007, India
REFERENCE 3 (bases 1 to 20310)
AUTHORS Goel,A.K., Rajagopal,L., Nagesh,N. and Sontli,R.V.
TITLE Direct Submission
JOURNAL Submitted (01-AUG-2001) C.C.M.B., Habshiguda, Hyderabad, A.P. 500
        007, India

REMARK 007, India
REFERENCE 4 (bases 1 to 20310)
AUTHORS Goel,A.K., Rajagopal,L., Nagesh,N. and Sontli,R.V.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-2002) C.C.M.B., Habshiguda, Hyderabad, A.P. 500
        007, India
REMARK Sequence update by submitter
COMMENT On Aug 1, 2001 this sequence version replaced gi:11693112.
FEATURES
            location/Qualifiers
            1..20310
            /organism="Xanthomonas oryzae pv. oryzae"
            /db_xref="taxon:64187"
            /note="pathovar: oryzae"
            218..1999
            /note="orf1; similar to ytfm from Escherichia coli"
            /codon_start=1
            /evidence=not_experimental
            /transl_table=1
            /product="putative outer membrane protein"
            /protein_id="AAG38834.1"
            /db_xref="GI:11693113"
            /translation="MLKVAFLFLLCTLEVPPLQAVAAQATIDKIEIKGLDDGDAEMIE
            MLOVSLIYEAVGKEQGESRLLEYLAQAEQRQREALEPPEGYSPITTLAAPRAGDVT
            VITVDRCGPVRVTRTSHISITGAEQDDRYLGODLKEPREGQVFSHPQYEAASKVIT
            RLRAEGRYFDAPETORRVAITTRAENHADIDILMWDSGRRYDMKVRVDYKFLGLEDP
            LVYWDGSGYVHEGKIDRLRESLTKLDYESTIDOPREPADDOGRPVYKFLTRAKT
            VYTACLSYSESAGYRGVGRYRYNARAKHKMDTQIDVAKNRSLSITRYKAPRMD
            GNTTASRLITDEQDITIDLRNVKLTGSRQINERNAIASINALERNRFSGGDFE
            GAUVETSTIYLPQDQANYINVDRLPERRGVSGOMIRGAGAGAGSDTNEFGVYGLR
            WELGADNSRLILREGGTTWTSDIAMPSPSLRFEAGSNSIRGYAFREVGRPTPRP
            KFLGAKNNVTASAEYEYHLNGSPMGCAVFNVSASFDDRPDMHTGIGRLMRSYVG
            PVRVDAHNLNPDQFOQLYIDIGANL"
            2299..5853
            /note="orf2; similar to ytfm from Escherichia coli"
            /codon_start=1
            /evidence=not_experimental
            /transl_table=1
            /product="putative periplasmic protein"
            /protein_id="AAG38835.1"
            /db_xref="GI:11693114"
            /translation="MKPOLRLGLGTLQNLNANVSNATLNLSKSEPTLPSPESLP
            QTTVPALDADKIDVQNLRTIDLOQPMYVLHVQGGSDVATGALTRDVIDTWDGF
            RLDGIVPSDDYHADLTATVLPAAKRTPASGLVARGDLMKEVALAKTRDVLADTDGDF
            SLVFEGRDPTWFAKAVTEALDTSLLIPAEQANPPATPIALNLQASGSKSNADLHG
            IKQGFASATIDPSPHIALADQVLTVEPIVITDFEGRTOGRPADLRDQNPSEFRVAV
            ASGLRFTPADPATPDAPLVPVELDARLVGAGTLKAMATIGRATVVERDGOOAEVFP
            SNGNDORQOLKOVAKTPTGGTLDLSGEVAMAPELQMDVNAOLAKTRDPTFAFGNGLN
            SGIASKGQLPAPAGCVSPFEATAEVPSLTGOLKORHLSNGKFRALRGEGESELQ
            LALGNSRINAKGLDQDLIAQQLDPLQDVLPGATGVNRQQLQVSGKRDAPDITAD
            IAGNLRMDTYSQOSTISLRRLPMRGSQDQALQGTALIEAGVLDVDSVRQARGAVEAL
            RLUDADIANMSASVALQGDVNRNGERMOCQVATLRIPASGDMALROPAQFSTGDAF
            TLDSTCLGAATGALCASAMPRAGMVGADLPISLVOPLMRKDDGQOITLYRGSLTF
            DGSEKPRGNAMGSGSLRASPFGCITGRTRAAVAGNPNRGLPRLDPPSMOADDTQO
            QIOGKLTGFCGAGVDAKPVNTGMAVYAPLNGELYLNNRSLWLELVADIVRPGRY
            EGHVSLKTRDKPPLGSDATISDFTAEVPSKGLALSEKGRDALPDSASITASAKS
            GPGTTLVDGSLWSEFTSPPLNLINIRGVNDLAVYNSELRIINPDMQFQITNNMILFG
            KGGISGQMOIRARPGREMIANGGLDVGGRKAYGQDILITRGOLTNMHNISYDPRVS
            LRAERKIDGITAGIDVGRASPRADVSEPMASQSELSLTVLGRGSLTYSDDTQO
            VSASAAALSASSLIASOTGATLGIDEGVSOSSLTGVSVGKTLSPKLYVGGVSM
            IGGSVLTILKTLKSRFDIEASSTVERKGSVMNRK"
            complement(5988..6934)
            /note="tnpX; nonfunctional putative transposase due to
            frameshift mutation"
            complement(7188..7625)
            /note="orf3"
            /codon_start=1
            /evidence=not_experimental
            /transl_table=1
            /product="hypothetical protein"
            /protein_id="AAG38836.2"
            /db_xref="GI:19483794"

misc_feature
            184..3539
            /note="Xanthomonadin biosynthetic gene"

CDS
            1..20310
            /note="Xanthomonadin biosynthetic gene"

```

CDS
/translation="MGRDAIATLIPHOSSMCIWEDIENDAOIRVLRSHNRSDAP
LRSNGRLALHLCXEYGAQAMAVGGLGASGRPRGMILVALRGVELHVTHTLPE
AIDCAOVLMOGEDSOQSEFLTHGEQLLEGRATVMLGAQA"
complement(7612..10035)
/function="probably involved in transport of xanthomonadin
to the outer membrane"
/note="Orf4"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative inner membrane protein"
/protein_id="AG38837.1"
/db_xref="GI:11693116"
/translation="MPRLRACVRLPQODTHWALFLNANRRIGLALLLAVAGF
MLSETLKVTDLRKFWPAPRTPAOKLLIELEGPGSRLLMLNSDPYTLAEOSQ
LOKALAAPELFEIVNGNGNAGIEAIKRLRLYLLSDSPKTPDAPLAEALDAR
VODLGRAPAAWYEPRLPRPTMELIHLAETLQPAAMPQTRNDYFPRACISALLIYR
RAAGFDPTQQLAYDAVQAFAKVSIGSTRLLTGPAAVEITARTQGEAWIGTL
DTVGLVLLLVAVRSWKIPVLGVLPLASAGLAVAVLFDVHGIVAFGFTLLGV
VODYPIHLSHQRPGLDPRENAHMLPTLVTVGVSTCIAYTFLEFGVDGLAVET
IGLATAVATTRMLPALIDPAPRDYADSRLLAVMRGVARLPRLPRLVLSAVAVSI
AVIATAGOFMONDSKLTPVPRKALADOMHILROELGAPDVRVLAIPANDAAVIA
ESDLPRLDAIVRNGDLGIDMAARILPSAKTOHARQALPDAQAQATETVATTP
FSDAFAPFLQDLDAKRAAPLIPKDTGSPLATSVGGLLGGDRSTALVSLTGLRD
PAVLAAYVSGAQLDLKDASESLAAVARGVLAALAVTAVIALSPRI
VVLPLMALTVLIALIRGTVELMLFHLIALIAGGLDVALFEDHAGDHAOL
RTLHALIVCSLMTLLVLFALLAASSIPVLRAIGSTVALGVLFNFILALVSRPALERT
ANGEPHAG"
complement(10120..10767)
/note="Orf5"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="AG38838.1"
/db_xref="GI:11693117"
/translation="MPMCRLLPMMLILLESAPLCAAPPETLDVGOVLERLARPAVS
TEFVELRGSSLLKTPLRVGHYRRPDQTLREVSVAPYHETTLTLDDEGLREGKSP
RRFSLSRVELAGLSGFGALLSGDRALAEONHVSAGGPOAMLLTDPKDAVVKO
VRQLRIGRKDELRCIESVPARKDVORTLLAGARPAAGVTAATATTLTCHSGA"
complement(10691..11665)
/note="Orf6"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative acyl transferase"
/protein_id="AG38839.1"
/db_xref="GI:11693118"
/translation="MSKHWKQREGGGRPALMLIRGINHGRGVNRRALYPTLYFL
LVRAPEERASRAVILARVLGRPATLRDVAARIHTFASTILDRVMLGCOMQRFVDSIG
LDQLHTQMDRGVLIFFSGHLSFDALRVLATERPVDQKVVVDKAKHNRMTDLGAL
NFOLANIIDACMDATSIYMAIKRATDAGALVALLLDRREGPRALPMLVIGNANRP
TSPMLIAALKVPAVILAFGLYHGNHVELAFETGSSGLDVRPHQBRPMLAVLMRDVAA
RLHNTCYAPYMWNFYDFPWNTHADVPRPAVDATVAVORTAMRRTA"
complement(11662..11949)
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative dehydratase"
/protein_id="AG38840.1"
/db_xref="GI:11693119"
/translation="MDFVYHDPCLPGHFGKRPVVYGVVLDHVLQAVEALHGPRAA
ARLPQVKFVQPLPCQTSVTLLEGDPRRMRFRVQRADVVVLVSGELVAEAGT"
complement(11961..12710)
/note="Orf8"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative acetyl transferase"
/protein_id="AG38841.1"
/db_xref="GI:11693120"
/translation="MPLARPSRLMAVNNLQLAFTLAFTAGGIYVALALMAITRDA
DMLRMSWMAVPLEFHAGVTVIVEGRERVMSKNYLFVSNHOSIIDICALFYALPV

CDS
PLRFLKEEMLKVPFVNMVATATGMLFLDRHRRAGALVROVAALLROGKOLCFPE
GTRSRGALLPFPKGLQAAIDAGVOVVPVALDCCGVLPVDFLFRVPGIIVRIGT
PIAVTGDPADRDQTLTEQAKAVAMHAPRI"
complement(12764..13438)
/note="Orf9"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="hypothetical protein"
/protein_id="AG38842.2"
/db_xref="GI:119483795"
/translation="MGVLLAVASPLAHMANASHRPEVAYIAGLLVLMVLEPVAR
RAMAMALALLLGLVLMHSVPYAMVVAAPVYFTGVAMFEGSRSLRGHTPLITFI
VEGLYGGAGMVPTEQRYTRRLTAMALLCGTLNVLVGLCAEESGVTOGVAS
PLPISDRAFLFANLLOCYIGGFEIYLLRGWFPQRPYRNLPDFLRQMARLGPDF
WRDLR"
complement(13485..13883)
/note="Orf10"
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="putative acyl carrier protein"
/protein_id="AG38843.1"
/db_xref="GI:11693122"
/translation="MLGLAVATATYTTADAPRTSRSRPHCCFLILPPYDLMKSM
SOTAAERELAEVLVESLNEEDVOPADIDEAPLPLNTGLGIDSLDALELALISKRYG
FQLRSDENRIRIFASLRALSAHEANKTY"
CDS
Alignment Scores:
Pred. No.: 4,24e-12 Length: 20310
Score: 353.50 Matches: 153
Percent Similarity: 35.11% Conservative: 91
Best Local Similarity: 22.01% Mismatches: 260
Query Match: 7.48% Indels: 191
DB: 1 Gaps: 20
US-09-914-168-2 (1-919) x AV010120 (1-20310)
QY 235 ArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIle 254
|||||
DB 437 CGGCAGACGCCG---GAGCGCTGGAGCCGTTGGCTTATCTGCCCAACCATCAGCTG 493
QY 255 IleArgAsnSerIleGlyGlu---ValAspValIleIleHis--AspLeuGlyGluPro 272
|||||
DB 494 GCTGCGCCAGCGCGCGGAGACAGGTGACCGGTGATCAGCGTGCAGAGCGG 553
QY 273 ValYrIleAspTyrAlaGluValAlaGluValArgGlyGluGlyAlaAspAspLysAlaIhe 292
|||
DB 554 GTGGCGGTGCGCACCTGCATATTTGCATCACCGGCTGGCGGACGACGACCGTTATCTA 613
QY 293 ThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisIleGlyLysTyr 312
|||
DB 614 GCGCAGACCTCAAGCACTTGCAGCGCGGCGGACGAGCGAGTCTTAGTCATCCGAGTAC 673
QY 313 GluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspLysArg 332
|||
DB 674 GAACCGACGACGAGCGGATACCGCGCGGCTGGCGGAGCGGTACTTGCATGCCGAC 733
QY 333 TrpLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIle 352
|||
DB 734 TTTCCCAACCCCGCGTAGCATCAGCGCGCGGACGCGCGCGCATATCGATCTGAAT 793
QY 353 TyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspProLys 372
|||
DB 794 TGGACACGCGCGCGCTAGCAGATGGGCAAGTGCGCTTC----- 835
QY 373 ThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeuGluGln 392
|||
DB 836 -----GACTACGACACTTC-----CGCATGGCTCTTTCGACCGC 871
QY 393 LeuLeuThrValAsnMetGlyGluAlaLysLeuGlnAlaValAlaGluLeuSerAsn 412
|||
DB 872 CTGGTACTAGTGGGACGAGGCGAGCTACTACACGACGACGAGCTGGATCGGCTACGCGAG 931

QY	765	AlaArgAlaGlyIleSerIleValTyrSerPheGlyAspAsnAlaTyrIleSerAsnArg	784
Db	1550	GTGTATGGCCAGCGCGCTGGTTTGTGGGCGGTGGGACAAACAGCGGTTGATCTTGGCC	1609
QY	785	AlaHisGlnMetThrIleGlyIleGlnAlaGlyTyrIleTrpSerAspAsnIleHis	804
Db	1610	GGC-----CAAGCGGCGACACCACCTGGACCAAGCACTGGTGGCG	1648
QY	805	ValProTyrArgLeuArgPheHisAlaGlyIleAspGlnIleArgGlyTyrAlaHis	824
Db	1649	ATGGCGGCAAGCGCTGGCTTTTGGCGCGATGGTGGTGAACACATCTCCGGTTATGGCTTT	1708
QY	825	AspSerLeuSerProIleSerAspArg-----GlyTyrLeuThrGlyGlyValLeu	842
Db	1709	CGCGAAGTCGGCGCGCGACCGCCCAAGCGCGCAAGTTCCGGCGTGGGCGCAACAGACTG	1768
QY	843	AlaValGlyThrAlaGlnTyr---AspTrpTrpIleThrMetIleAspLeuArgLeuAlaVal	861
Db	1769	GTTACCGCGCATCGCGGATATGACAGACACTCTCAACGGTGCGCGCGCGCGCGGTG	1828
QY	862	PheGlyAspTrpIleGlyAsnAlaTatTyrAspLeuGlyPheThrHisAspTrpIleGlyAla	881
Db	1829	TTCGTGTACACAGCGCGACGGCTTCTGAGAC-----CGCGCGACCTGGCACACCGGATC	1882
QY	882	GlyValGlyValArgTyrTrpAlaSerProValGlyGlnValArgValAspValAlaThrGly	901
Db	1883	GGCTTCGGCTCGGTGGCTGGCGTGGCGCGGTGGCGGTGGCGGTGGCGGTGGCGGTGGCG	1942
QY	902	ValGlyGlnIleGlyAsnProIleGlyLeuHisPheThrPheIleGly	916
Db	1943	TTCAGCAGACCGCGACCGCGAGTTCTCAAGCTCTACATGATGATCGCG	1987
RESULT 22			
AE003957			
LOCUS	AE003957	10074 bp	DNA linear BCT 15-JUN-2001
DEFINITION	Xylella fastidiosa 94sc, section 103 of 229 of the complete genome		
ACCESSION	AE003957	AE003849	
VERSION	AE003957.1	GI:9106207	
KEYWORDS			
SOURCE			
ORGANISM	Xylella fastidiosa 94sc.		
	Xylella fastidiosa 94sc		
	Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;		
REFERENCE			
AUTHORS	1 (bases 1 to 10074)		
	Simpson,A.J., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M.,		
	Alvarenga,R., Alves,L.M., Arata,J.E., Bala,G.S., Baptista,C.S.,		
	Barros,M.H., Bonaccorsi,E.D., Bordin,S., Bove,J.M., Briones,M.R.,		
	Bueno,M.R., Camargo,A.A., Camargo,L.E., Carraro,D.M., Carier,H.,		
	Colauto,N.B., Colombo,C., Costa,F.F., Costa,M.C., Costa-Neto,C.M.,		
	Costa,L.H., Cristofani,M., Dias-Neto,E., Docena,C., El-Dorri,H.,		
	Fadoniani,A.P., Ferreira,A.J., Ferreira,V.C., Ferro,J.A.,		
	Fraga,J.S., Franca,S.C., Franco,M.C., Frohme,M., Futari,L.R.,		
	Garnier,M., Goldman,G.H., Goldman,M.H., Gomes,S.L., Gruber,A.,		
	Ho,P.L., Hohnsels,J.D., Junqueira,M.L., Kemper,E.L., Kitajima,J.P.		
	and Marinho,C.L.		
TITLE	The genome sequence of the plant pathogen Xylella fastidiosa. The		
	Xylella fastidiosa Consortium of the Organization for Nucleotide		
	Sequencing and Analysis		
JOURNAL	Nature 406 (6792), 151-157 (2000)		
MEDLINE	20365717		
PUBMED	10910347		
REFERENCE	2 (bases 1 to 10074)		
AUTHORS	Simpson,A.J.G., Reinach,F.C., Arruda,P., Abreu,F.A., Acencio,M.,		

Lambais,M.R., Leite,L.C.C., Lemos,E.G.M., Lemos,M.V.F., Lopes,S.A.,
Lopes,C.R., Machado,J.A., Machado,M.A., Madeira,A.M.B.N.,
Machado,M.F., Marino,C.L., Marques,M.V., Martins,E.A.L.,
Martins,E.M.F., Matsukuma,A.Y., Menck,C.F.M., Miracca,E.C.,
Miyaki,C.Y., Monteiro-Vitorello,C.B., Moon,D.H., Nagai,M.A.,
Nascimento,A.L.T.O., Netto,L.E.S., Nham Jr.,A., Nobrega,F.G.,
Nunes,L.R., Oliveira,M.A., de Oliveira,M.C., de Oliveira,R.C.,
Palmeri,D.A., Paris,A., Peixoto,B.R., Pereira,G.A.G., Pereira
Jr.,H.A., Pesquero,J.B., Quaggio,R.B., Roberto,P.G., Rodrigues,V.,
de M. Rosa,A.J., de Rosa Jr.,V.E., de Sa,R.G., Santelli,R.V.,
Sawasaki,H.E., da Silva,A.C.R., da Silva,F.R., da Silva,A.M., Silva
Jr.,W.A., da Silveira,J.F., Silvestri,M.L.Z., Siqueira,W.J., de
Souza,A.A., de Souza,A.P., Terenzi,M.F., Truffi,D., Tsai,S.M.,
Tsunako,M.H., Vallada,H., Van Sluys,M.A., Verjovski-Almeida,S.,
Vettore,A.L., Zago,M.A., Zatz,M., Maidanis,J. and Setubal,J.C.
Direct Submission
Submitted (02-JUN-2000) Organization for Nucleotide Sequencing and
Analysis, Bioinformatics Lab - IC/Unicamp, C.P. 6176, Campinas, SP
13083-970, Brazil
FEATURES
source
location/Qualifiers
1..10074
/organism="Xylella fastidiosa 9a5c"
/db_xref="taxon:160492"
/clone="9a5c"
complement(271..759)
/gene="Xf1230"
complement(271..759)
/gene="Xf1230"
/note="similar to SPIP76270 (percent identity: 49 %/query
alignment coverage: 87.7 %/subject alignment coverage:
77.6 %); identified by sequence similarity; putative; ORF
located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAFP8404.1"
/db_xref="GI:9106208"
/translation="MPLAPPIITGNKPEYQALQAKALIHGETDRIANANALAL
IFHSPLGLMWGEFYDGRLELVGPFQGLPACIRIPIDKVGCGTAARTRRORIDVD
TFPDHACDSASRSEIVVPLIKDPTLIGVLDLSDPTCARFSDQGTGIESLARFVSS
LR"
1052..2905
/gene="Xf1231"
1052..2905
/gene="Xf1231"
/note="similar to SPIP39320 (percent identity: 27 %/query
alignment coverage: 95.9 %/subject alignment coverage:
102.6 %); identified by sequence similarity; putative; ORF
located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAFP8404.1"
/db_xref="GI:9106209"
/translation="MKSRCFWNVRFITLSMLPRYTAQMLRRIIPMLISMLFASLASLA
QAVGVKIQIKGDKGDVEMKMNITVSLSYVCKEKEFEFLVLLSAEQTRO
ALPEFGYAPTRIDARONDHITVIVYDVGEPVRAQAVAMTGAADHAYDRL
EDRPRKIGELFHNHYEASKYRITRRLAERGVFPADFTHREVTAAAHADIDLIME
SGRYDMGVPVRFHYDFHGLFNPLVYWDGSGYKIDRLRESLTKLPPESIDIO
PKREBAPBENVPYDVAKLERAKSKITTAGISYSGESGCLAGACVERRMARGRKMA
RLDYAOMKSLTAYQIPAFRWLDGWTIFARAYADEQTKYIDLNRVKISARSGQINR
HLTATASLNLKGRMRYRADDTNTVAYQOSTLVYPOLEASYVDVADATFPRNSAAT
VLTGGSALGSKSNFTQHLGQLFGLGASRLILLEGAGTWTASDLVAMPISLR
FAGCVANSIRGAYREVEGVPRTAKPDAPALGAKHVSAGAEFEYTKGSGPFGGAVVDGS
SAFNRYPMHMTGIGILRYSRSPGVAVDLARGLNSDSKIQTLIDIGANL"
3184..6723
/gene="Xf1232"
3184..6723
/gene="Xf1232"
/note="hypothetical protein; identified by sequence
similarity; putative; ORF located using
Glimmer/RBSfinder/Start codon shift: 411"
/codon_start=1

/transl_table=11
/product="hypothetical protein"
/protein_id="AAFP8404.1"
/db_xref="GI:9106210"
/translation="MDPDIWLLRLRLDYLEMNSASLSGSEEPNFPSPDVL
SINPLAQLDITTDINFRISLKQPLHLRIKGAVANVADDELQRLAVSDRNI
RADGVYRRQDSTDLTATVAVPSRGRPASGVLTRDLDRLQALAGRPAQQA
DVLGSRAPRMTFSAASEHFFGLLPCELTRTSATLALNMOAGCGGADLQQA
FKQGGOKVLPDQSLRLANOMMPEVITDAGGRTNLRKGYDLDPARTRFPSNA
RGLSPSPERDNTQVRAVPVYLKAGVSGSTLTANMAAGARLTRDQARLFDV
RQDAHATFKQVQATPGCLDFTGVTWAPLRNDLAKELQDFGFTPGMWVS
GOLTSQGRQRTVYSSGIEALVAVPGLNGELRQPMQASGFTFQDQEGEAFQIGDS
HISAGKGPHLDLAVQOSLNLNDLPSKEGTLVLRLRPARTPVYIVLNGSL
RWNDRARQFALHGLPMRNGKTLDAOEIEIGTLDSVIRHSGAVENLQADV
RHAMLLHLHGVRHDGARMWEDLSQLEPPEKGPAMALRAPOEVYMAVPFTLSATC
LDAOGGALCLHAPDHGICLVKGOALPTLYVPLPKNGRSINLRGALTJDCIHRP
HDAWEGSFRRASSLEGTRIGRPRADVLGVQDRELYRYDQPSVOVDTLSHIVKRL
GIFQNGFVDAKEMTGMDDPYAPPLGDLVYNNRSRYLWLEMPDIGNPGLLEGHVS
RGTSGPILGGEATLRDFTAEYPSMGLNLSOGGRTAQPDGSARILASVRSGDTLN
IDGGLSMFGENNPIQLNIRGNVAVLNTSELIVSDPLRGLINTNMOLQGEAVSE
ADIDLERDRGVSEEDVYVLDPPVPERTAVSPLEMNLRIVGDKVYNNMGFGLKGLS
GOMOTRARPGRHMTANGGLDVGRYKAYAGODLTITRGOLVSMNNTVSPRRVITRARR
IADITRAGIDVSGRAQSPRYQWMSDVMSQSESLTVLGRKLTGTSASDENOQVASAA
ALSAGGLIASOLGAKLGLDDAGISQSTLGGVSGFGRYTLSPKLYGVGVSMSGSS
VLTLYLLRGRFDLEVSSTVETKGSVMRKER"
complement(6826..6945)
/gene="Xf1233"
complement(6826..6945)
/gene="Xf1233"
/note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAFP8404.1"
/db_xref="GI:9106211"
/translation="WLFPEHEHIGNSAMSLAIGASFLVIDPQFTGQANRCP"
7500..8396
/gene="Xf1234"
7500..8396
/gene="Xf1234"
/note="similar to SPIP7541 (percent identity: 62 %/query
alignment coverage: 98.3 %/subject alignment coverage:
99.0 %); identified by sequence similarity; putative; ORF
located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="carboxyphosphonoenolpyruvate phosphonmutase"
/protein_id="AAFP8404.1"
/db_xref="GI:9106212"
/translation="MSTSLPIQTPGOCFPTALMTERPLOVIGAMNAYHALLAORAGYR
ALYLSGGVANGSGICLPIDGMNTLEDVLRIRTDVDCPLVINDPQFGSAGNIE
RIKTLIKAGAAACHTEDQVAKKRGCHRGKATVLSLDMADYVKAADRTAAFLFI
ARTDAIQAEVESAIERAOAYVEAGADSTFAVAVSDLETYRHFVERIRVPLANNTEF
LYERLDYHAFEGRLDALARSR"
complement(8703..8778)
/gene="crna-Lys-1"
complement(8703..8778)
/gene="crna-Lys-1"
/product="crna-Lys-1"
/note="Found by tRNAScan"
9291..9446
/gene="Xf1235"
9291..9446
/gene="Xf1235"
/note="hypothetical protein; identified by sequence
similarity; putative; ORF located using Glimmer/RBSfinder"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAFP8404.1"
/db_xref="GI:9106213"

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 201050)
AUTHORS Boucher,C.A.
TITLE Direct Submission
JOURNAL Submitted (05-DEC-2001) Genoscope and CNRS UMR-8030, 2 rue Gaston
Cremieux, CP5706, 91057 Evry Cedex, France, Laboratoire de Biologie
Moléculaire des Interactions Plantes-Microorganismes INRA-CNRS,
BP27, 31326 Castanet-Tolosan Cedex, France, Fondation Jean
Dausset-CEPH, 27 rue Juliette Dedu, 75010 Paris, France, LMGM CNRS
118 Route de Narbonne, F 31062 Toulouse Cedex 4, Genoscope and INRA
URCV, 2 rue Gaston Cremieux, CP5706, 91057 Evry Cedex, France,
Laboratoire de Biometrie et Intelligence Artificielle INRA, BP27,
F3136 Castanet-Tolosan Cedex, Laboratoire de Genetique Cellulaire
INRA, BP27, F31326 Castanet-Tolosan Cedex
Christian.Boucher@toulouse.inra.fr
http://sequence.toulouse.inra.fr/R.solanacearum.html.

FEATURES
source Location/Qualifiers
1.201050
/organism="Ralstonia solanacearum"
/strain="GM11000"
/db_xref="taxon:305"
206..1447
/gene="RSC1326"
/note="RS02848"
206..1447
/gene="RSC1326"
/EC_number="2.6.1.1"
/function="small molecule metabolism; amino acid
biosynthesis"
/note="Product confidence : probable
Gene name confidence : hypothetical
Predicted by Codon_usage
Predicted by Homology
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE ASPARTATE AMINOTRANSFERASE PROTEIN"
/protein_id="CAD15028.1"
/db_xref="GI:17428341"
/translation="MKTIQSAKLNVCYDIRKPVLEKAKOMESEGHQITIKLNIGNLA
PGCFAPPEIIOODMIRNLPNSAGYSDSGIPAPKAKVMHYTOOGIKNTLIDYIYN
GASSELLIATNALDLADGDELLPAPYPLMTAMTSLSGTPVYILCDSENGMRPDD
IRAKTPTKGTIVINPNPRTGALISDALHDIATAREGLVIFADEYDKYLFUDN
RHTAASUSEDEVLTFTFNSLSYSCGIRACGMVVSQKRPADYIEGLNLSMRL
CANTACGQWAIQTALGYSINDLVAPCGMRRLAYELITAIPEGVSCVKKPAALYM
FPRLDSPYIDDQTFIRQLLEERVLVVGCTGFNMHSPDHFRIVFLPHEDDLREAI
GRIARFLERYRQRHGTGIRKA"
1538..2857
/gene="RSC1327"
/EC_number="1.1.1.3"
/function="small molecule metabolism; amino acid
biosynthesis; threonine"
/note="Product confidence : probable
Gene name confidence : hypothetical
Predicted by Codon_usage
Predicted by Homology
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE HOMOSERINE DEHYDROGENASE OXIDOREDUCTASE
PROTEIN"
/protein_id="CAD15029.1"
/db_xref="GI:17428342"
/translation="MKPIKIGLLGTVGGGTLKVLQNRQEIIRRRAGGIEVALIAV
RNVERRAIVQETGASVPVDDPAVVDSPEDIDIVVELIGDYDLTRELVLRAIANKH
VYTKAKALLAHGNEIFKAADQKGVAVFAEAVAGGIPILKALREGTLRIEMLAGI
INGTNPFLISEMRKDGDTGLVLAQORIGVGAEDPTDIEGFDAAHKLTLSALAFG
VPVQERAVESITKLEAVNDIRYAEELGRIKLLGLTRAEAGIELRHPTLVPRTRL

gene
CDS
2889..4334
/gene="thrc"
/note="RSC1328; RS02850"
2889..4334
/gene="thrc"
/EC_number="4.2.99.2"
/function="small molecule metabolism; amino acid
biosynthesis; threonine"
/note="Product confidence : probable
Gene name confidence : probable
Predicted by Codon_usage
Predicted by Homology
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PROBABLE THREONINE SYNTHASE PROTEIN"
/protein_id="CAD15030.1"
/db_xref="GI:17428343"
/translation="MQRSTRGHAEPQSFRIILGILAPDGGYLLPPOYQVSAEELT
RMRDLPYADIAFEILRKFAVDIPEDDRLALRTTYTEVCNARNGONTADITPLHL
GEGAGTSLDLCNCPYLAQKDMAMOLNLNLEAREHAEINLIGATSGDGSAA
EYAMRGKRGYRVFMLSHPKMSAFQTAQMSLODKNFNIAIEGFPDADIDYAAVN
DLEAKRQKIGTVNSINMARVAVQVYFFGYLLAARKIGDKVSCVPSGNGVWCG
HIAKMGDLPIDKLIVATNENDYIDDEFRTGTGYRVSAAQGYHNSSPMDISKSNFR
FVFDLGRDPAKLAQLFRDVERGGEFMSKPEPDRIAEGFVSGRSTHEDRYTID
VPRYGYMVTHTADGVKVARENLVPPVAVVLETTALPAKFGDTIRALDERPRPAG
YADIESLPQRFEPWADADRIKGYIAQHTGL"
4361..4957
/gene="RSC1329"
/note="RS02851"
4361..4957
/gene="RSC1329"
/function="miscellaneous; hypothetical; global homology"
/note="Product confidence : hypothetical
Gene name confidence : hypothetical
Predicted by Codon_usage
Predicted by Homology
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAD15031.1"
/db_xref="GI:17428344"
/translation="MAEAPAFSCULDIRACLCEGALPHEPPVVLADRARVILVQ
APGRIVHETGIPWNDSGRRLREMLRIGYDAFADRIAVPMPGFCPPGDPNGDLPP
RPACAPQWHRPRLTAMPDCLKLALVGYAORYLVDPGRGMLTEVVRARGVRDAEGRV
LFLPLPSPRPNLAMFKNPWFDAQIVRLRAALAAAFV"
5037..6296
/gene="moea2"
/note="RSC1330; RS02852"
5037..6296
/gene="moea2"
/function="small molecule metabolism; biosynthesis of
cofactors, carriers; molybdopterin"
/note="Product confidence : putative
Gene name confidence : putative
Predicted by Codon_usage
Predicted by Homology
/codon_start=1
/evidence=not_experimental
/transl_table=11
/product="PUTATIVE MOLYBDOPTERIN BIOSYNTHESIS MOEA
PROTEIN"
/protein_id="CAD15032.1"
/db_xref="GI:17428345"
/translation="MSANDAPAYKPSLITLLEALLDOLLAARGLTETETVPTTHANGR
VLAQAVRSGLVNPPADNTQMDGYAVRADITATGTRLKVAQRIPACOVGRPLDAGEAA

OY	366	LysATGgIleuleuLeuGLI-----GIHeuLeuthrValAsnMetGLyglAlatry---	402
Db	101298	GAGGGCGACTCTCCGGCAGCATGCAGTGTGCCACGGCCAACTGGCTGTGTGGTAACC	101357
OY	403	-----AsnLeuGlInAlaValArGALeUsenrAsnSpleuILealArhArgTyPhe	420
Db	101358	AAGAAGCACTCTACTCCCAACAGAGTCGAGCGCCAGCTGAAGACGCTGGCTGTTCC	101417
OY	421	AsnMetValAsnThhGluIlleValPherosLIuArgLIuclInllegInasnsAgLIuVal	440
Db	101418	TACCTGGACCCCCGCTACTGTGAATTCT-----GCCATCGAATGACCCAGTG	101455
OY	441	SerPhelugInserSerSerSerArgThrGluProAlagInValAspArgSerThrlEu	460
Db	101466	TGCATCACCGCGGACAGAACAGACATCTACTGACGGTAGAACATCCACAGAGGTGAG--	101522
OY	461	GIUpProvalIIlegIuThrValIGIleuThrAspLIyleUmetAspIIeserProIle	480
Db	101523	CAGTCAAAGCTGTCCGACATCAACTGCACCGCGAGGTGCTGTCTCAAGCAGGCCGAGATG	101588
OY	481	GIU-----PheSerLaseAsnLeu-----	487
Db	101583	GAGAACTGATCAAGCTTAACACAGGTGACGTGTTCTGTCGGCCAAGCTGTGGCCACC	101642
OY	488	-----IlegInAspLyIsleuAsnLeuValAlaAlaLySaLaArgHISLeuTyraSP	504
Db	101643	ACCAGCGGATCACCGACTCTGTGGCAGCAGTACCGCTACCGCTTCGGCAGCATCAACCG	101702
OY	505	MeTrprosp-----AspArgValLeuAlaIleasn-----HisAspAsgly	518
Db	101703	CAGCGCAGATCAACCAAGACAGCGACGGTGTGCTGCATCAGCTGGTGTGATCCGGGG	101762
OY	519	-----ValAsnArg---SerIIleuGlyArgIleserAspAlaValaserAla	533
Db	101763	CGCGCGTGTACGTAGCGCCGGGTGACAGGTGTGGCAAGCAAGACGGCGAGCAAGTG	101822
OY	534	ValAlaAlargAlaIlleleuproAspLIuser-----GluasnGluValIIleaspLeu	550
Db	101823	GTGCGCCCGCAGATGCGCCAGATGGAGTGTGTGTGTTCAAGCGCGAAGCTGCACCTTC	101882
OY	551	PROGUARG-----ThralalaValAsnArglys	560
Db	101883	TGCGAGAACCGCATCAACCGGACGGGCTACTTCAACGACAACCAATCACCCAGAAAC	101942
OY	561	ThrProAlaAspValTyrgInSerLyIsylValProleuTyrgValPheValAlaserASP	580
Db	101943	GTGCGGGGCG-----ATGTCCGACAGGTGGATGTCTCAACGTAACCTGTG---ACCGAG	101990
OY	581	LYSProArgAspGlyGlnIIlegLyLeuGlyTyrrplySerAsprThrclyuThrarArgLeuVal	600
Db	101991	AAGCGG----ACCGGCSKATACAGCTGGCGCTGGCTTCTTCGTGACSSGACAAAGCTGGTG	102047
OY	601	Thrlus-----PheLIuHISAsnLeuIIleasnArgdsrGlyTyrgInAlagIlaGlylaG	618
Db	102048	CTGTGCGCCCGTATTCGCGAGACAAAGATGTGTTCGCGTCCGGGACACAGCTGGCGTGGAC	102107
OY	619	LeuArgLeuSerGluAspLyIsylValLyIsleuTyrrAlathrltyrProleuseSerHIS	638
Db	102108	GTGAACACTCTCAAGTCCAAACCGGACACCATCTCCGTGCAC-GCAGTASACCCGTAATTCTAC	102166
OY	639	-----ProleuAsnAspGlinleuTyrgAlathrltyrLeuGlyTyrgInGlnGlu	653
Db	102167	GTCTGATGGCTATCAACCCGCTGCAC-----CGACCTGTACTACCGCACAGTACCGCCC	102217
OY	654	ValPheclyllISerThrasnGlyPheAspLeuSerThrArgThrlleuGlnsgIuile	673
Db	102218	GCTGTACTACACCGCGACACAGACTACCGCATCTGCACAGAGCGGCAAACTCAAGTT	102277
OY	674	SerArgserIleIIlegInasnclyGlyTyrrAsnArgThrltyrserLeuArgTyrrArgLeu	693
Db	102278	CGGCGTGGCTTTCTC-----GAAACCGAATAACCGTGTCTTTCGGCATTCGGGCTA	102325

QY	694	-----AspLyseuLeuys--ThglnalaproPrognUthrPrpGlnaspLeuProva	710
Db 102326	CGAGCGCACGACCATCGACGTACCTCCACACGCCGTTGGTCTACACAAATACCTGGCC	102385	
QY	710	LasphvehalasnGlyLysProSerGlnGlnAlaLeuLeuAlaGlyValAlaValHisLy	730
Db 102386	CAGG-----AACGGCCGCGATTTCGACAACTCCCGATCACCC--ATCGCGTGTGCGA	102436	
QY	730	gthrValAlaAspAsnLeuValAsnProMetArgGlyIlyrArgGlnArgIlyrSerLeuGl	750
Db 102437	GGACCAAGCGCGACACGCCCTGTGTGCTCCAGCGCGGCGCGCTACACAGCGCCACCTyGA	102496	
QY	750	uValLyseSerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIlyLeSe	770
Db 102497	GTTCCGGCATTCGGGCG-----GGCGACCTCGAGTACATCAACGCTGGCTACTACCA	102544	
QY	770	rcIlyValIlyrSerPheGlyAspAsnAlaTyGlySerAsnArgAlaHisGlnMetHcGl	790
Db 102545	GCACCAAGTAC-----TTCTATTCCGCTGTGCGAAGTGGTTCACGCGCGCTT	102589	
QY	790	yGlyIlyleGlnAlaGlyIlyrIleTyrSerAspAsnPheAsnHisValProTyArgLeuAr	810
Db 102590	CACACACAGCATGGCGCTACGGCGCCACGGCTACGGCACACAGACCTCCGGGTCTTAAGA	102649	
QY	810	gPhePheAlaGlyIlyAspGlnSerIleArgIlyTyAlaHisAspSerLeuSerProIl	830
Db 102650	CTACATACCGCGCGGCTATACGCTCGCTCGCGCGGTACAGACACACAGCTGGCGCGCG	102709	
QY	830	eSerAspLyseGlyTyIlyrLeuThrGlyGlyGlnValLeuAlaValIlyrThrAlaGlyIlyAs	850
Db 102710	CGACGCCAACAGCGCGTGGCCATGTGTGGCGCCACGACAACTTCGTGGCAACATGGATTCAT	102769	
QY	850	nTyrgIlyPheMet-----LysAspLeuArgLeuAlaValPheGlyAspIl	865
Db 102770	CTTCCCGCGTGGCGGCTCGGCGGTGAGACCGCACCTCCGCTTCACCTCTTCCTGACATA	102829	
QY	865	eGlyAsnAlaTyArgAspLyGly-----PheThrAsnSprThrLySIlleGlyAlaGl	882
Db 102830	CGGCAACGCTGTTCGCCGGAAGGCGACCCGCTCAAGGTCCGCGCATGCGCTACGCGCGG	102889	
QY	882	yValGlyValArgTyrPalasSerProValGlyGlnValArgValAspValAlaThrGlyVal	902
Db 102890	TTTTGGCGCTGTGCGTGTGCGCGATCGCGTCTGAAGATCACATGGGCTCCGAT	102949	
QY	902	llySerIly-----GluGlyAsnProIleLyLeuHisPhePheIleGlyThrProPhe	919
Db 102950	CAGAAGGAAGCGCGAAGACCAAGCAGACCTTCAGATTCAGATCGGCGGACTGCATTTC	103007	
RESULT 25			
AX155443		2379 bp	Linear
AX155443			
DEFINITION	Sequence 6 from Patent WO0138350.		
ACCESSION	AX155443		
VERSION	AX155443.1 GI:14536771		
KEYWORDS			
SOURCE	Neisseria meningitidis.		
ORGANISM	Neisseria meningitidis		
	Bacteria; proteobacteria; beta subdivision; Neisseriaceae;		
	Neisseria.		
REFERENCE	1 (bases 1 to 2379)		
AUTHORS	Giuliani,M.M., Pizsa,M., Rappuoli,R. and Holst,J.		
TITLE	85kda neisserial antigen		
JOURNAL	Patent: WO 0138350-A 6 31-MAY-2001;		
	Chiron Spa (IT) ; Statens Institut for Folkehelsetse (NO)		
FEATURES			
Source	Location/Qualifiers		
	1..2379		
	/organism="Neisseria meningitidis"		
	/db_xref="taxon:487"		
BASE COUNT	636 a 755 c 596 g 392 t		
ORIGIN			
Alignment Scores:	8.84e-08	Length:	2379

Score: 265.50 Matches: 195
 Percent Similarity: 34.28% Conservative: 120
 Best Local Similarity: 21.22% Mismatches: 348
 Query Match: 5.62% Indels: 257
 DB: 6 Gaps: 39

US-09-914-168-2 (1-919) x AX155443 (1-2379)

QY 116 ThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
 DB 46 TCCGCTTTGGCATTGGCCGACTTCACCATCCAGATCCGTGCGAAGGCTTGACGGCT 105
 QY 136 AsnAsp-----TyrlleProGluTyrGlnGlyGluGluProAsnSer 149
 DB 106 ACCGAGCCGAGCAGCATTCATCACTACCTGCCCTCAAAAGTCGGGAGCACTACACAC--- 162
 QY 150 GluValValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
 DB 163 -----GACACACGCGGCACTGCCATCATCAAAAGCCTGTAC 198
 QY 170 AlaArgLeuPheAsnAspGlyValAlaAsnLysValPro-----ArgLeuLysAlaLys 186
 DB 199 GCCACCCGTTCTTGTGACGACGTACGAGTCAAACTCGCGAGCGGCACTTCTGCTGACC 258
 QY 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
 DB 259 GTTATCGAACGCCGCCACCATGCTCGCTCAACATCACCGCGCCCAAAATGCTGCACAAAC 318
 QY 207 GluProTyrAlaAsnIleLysAlaAlaLeuGluAsp-----IleThrGlnGluSerAla 224
 DB 319 GACGCC-----ATCAAGAAAAACCTCGAATCGTTCGGGCTGGCGACGTCCGCAATAC 369
 QY 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
 DB 370 TTTAATACGCGACACTCAACACGACGTCCCGGCTGAAAGAAATACCTCGCGGCT 429
 QY 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
 DB 430 GCCAAACTCAATATCAAAATCAAGCCCAAGTAACCAAAATCGCCCGGCAACCGGCTGCAC 489
 QY 250 IleAspLeuSerIleIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
 DB 490 ATCGACATCAAGATGTGACGAGGCAAAATCCGCCAAAAATCACGACATCGAATTGCAA--- 546
 QY 270 GlyGluProValTyrIleAspTyrArgAlaValAlaGluValArgGlyGluGlyAlaAspAsp 289
 DB 547 GGCACCAAGTCTATTCCGACCGCAAACTGATCGG-----CAG 585
 QY 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisHis 309
 DB 586 ATGTGCTGACCGAAGCGGCAATTGGACATGGCTGACACGAGGACCGGTTGCAACCGC 645
 QY 310 GlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPhe 329
 DB 646 CAGAAATTCGCCCAAGCATGGAAGTAACGCACTTCTACGAAACAGCGCTACTTC 705
 QY 330 AspGlyArgThrLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspVal 349
 DB 706 GATTTCGCTATCTCTGATACCGACATCCAAACCAAGCAAGCAAAACCAAGCAGCAACCTTC 765
 QY 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIle 369
 DB 766 AAATATCCCTCCACGAGGCGGAGGATTCGCTGCGGCAAGATG-----TCGATT 816
 QY 370 AspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeu 389
 DB 817 GAAGGCGACACCAACGAA-----GTCGCCCAAGGCGCGAA 849
 QY 390 LeuGluGlnLeuLeuThrValAsnMetGlyGluLysValTyrAsnLeuGluAlaValArgAla 409
 DB 850 CTGGAATAACTGCTGACCATCAAGCCCGCAAAATGTTACGAAGCCGACAGATGACCGCC 909
 QY 410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhe 429

DB 910 GTTTGGGT----- 918
 QY 430 ProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArg 449
 DB 919 -----GAGATTCTGAACCGCATGGGCTGGCAGGCTACCATACAC----- 960
 QY 450 ThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeu 469
 DB 961 -----GAAATCAGCTACAGCCCTGCGCAACCCGCAACCA 999
 QY 470 ThrAspGlyIleLeuMetAspIleSerPro-----IleGlu 481
 DB 1000 ACCGTGATTTGCTGCTGACATCGAACCGGCGGCAAAATGTACGTCAACGAATCCAC 1059
 QY 482 PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHis 501
 DB 1060 ATCACCAGGCAACAAACCCGCGACGAA-----GTGCTGGCGCCGCAATTTGCCCA 1113
 QY 502 LeuTyrAspMetProAspAsp-----ArgValLeuAla 512
 DB 1114 ATGGAATCGCGCCTTACGACACCTTCCAAAGCTGCACAAAGAGCGCTGACGCTT 1173
 QY 513 IleAsnHisAspAspGlyValAlaAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532
 DB 1174 TTGGGCTACTCTGACACGTACAG-----TTTGAT 1203
 QY 533 AlaValAlaAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
 DB 1204 GCCGTCCCGCTTGCCTGACGCGGCAACAA-----GTGCAATTGG----- 1242
 QY 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValPro 572
 DB 1243 AACATGAGGCTTACCCAGACGCTTCAC----- 1269
 QY 573 LeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeu-----Gly 590
 DB 1270 -----GGCTGCTGCACTTGAGCGCGGCG 1293
 QY 591 TrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArg 610
 DB 1294 TGGGTTCAGATGATCCGCG-----TTGCTCATGTCCGCGCGGTATCCGACGCAACCTG 1347
 QY 611 AspGlyTyrGlnAlaGlyValAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeu 630
 DB 1348 TTCGTTACGGGGAAGTCGCGCCCTGCGCGCTCGCGAAGCAAA-----ACCACGCTC 1401
 QY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
 DB 1402 AACGCTGCTGCTGCTTACCGACCGCTACTTACAGGAGAGGCGGTGAGCTGGGCTAC 1461
 QY 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670
 DB 1462 -----GATATTTACGGAAACCTTTCGACCGCGCAAGCAAGATCGACGCTCAAAACA 1515
 QY 671 HisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThrTyrSerLeuArg 690
 DB 1516 TATTAACACCAACCGCC-----GGCGGCGGCTAAGATGGG-TATCCCGCT 1562
 QY 691 TyrArgLeu-AspLysLeuLysThrGlnAlaProProGlu-----ThrTyr 705
 DB 1563 TACCAATACGACCGCGTCAATTTGCGGCTGCGCGGAACACGTACCGTCAACACCTA 1622
 QY 705 PclAspLeuPro-----ValAspPheValAsn-----GlyLys---ProSerG 719
 DB 1623 CAACAAAGCACCAACAGCTATGCGCACTTTCAGAGAAATAGGCAAAACCGACGCGCGC 1682
 QY 719 nGluAlaLeuLeuAlaGlyValAlaValHisLysThrVal-----Al 733
 DB 1683 AGACGGCAGCTTCAAAAGGCTGCTGTACAAAGCAACCGCTGCGTGGGCGCAACAAGAC 1742
 QY 733 aAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySe 753

```

Db 1743 GCACAGCGCGTCATGGCCGACGGCGGCTAC----- 1773
Oy 753 rSerGlyLeuValSerAspAlaHisMetAlaIleAlaArgAlaGlyIleSerGlyValTy 773
Db 1774 -----CTGACCGCGGTAATGCGCAATGCGCGCGCGGACCAACTGCATATA 1823
Oy 773 rSerPheGlyAspAsnAlaTyGlySerAsnArgAlaHisGlnMetThrGly----- 790
Db 1824 C-----TACTCCGCCACCSCAACCAACCAACCGTTCCTCC 1859
Oy 791 -----GlyIleGlnAlaGlyTyrIle 797
Db 1860 CTTAAGCAAAACCTTCACCTGATGCTCGCGGCGGAAGTGGCATTTGGCGGCTAC-- 1917
Oy 797 eTTPSerAspAsnPheAsnHisValProTyrArgLeuArgPheAlaGlyIleAspGly 817
Db 1918 -----GGCAGACCCAAAGAAATCCCTTCTTGAACACTTACGGCGCGGCTGGG 1970
Oy 817 nSerIleArgIleTyrAlaHisAspSerLeuSerProIleSerAspIleGlyTyrLeuTh 837
Db 1971 TTTCGGTCCCGGCTACGAAAGCGGCGACGCTCGGCGCG-----AAAGTGTATGACGA 2021
Oy 837 rGlyGlyGlnValIleValIleGly-----ThrAlaGluTyrAsnTyrGluPhe 854
Db 2022 ATACGGCCGAAATATGACTACGGCGGCAACAAAGCCAAACGCTCCCGCGGCTGT 2081
Oy 854 r-----LysAspLeuArgLeuAlaValPheGlyAspIle 865
Db 2082 CTTCGGCATGCCGCGTGGCAAGACGACGACCGTCCGCTGAGGCTGTGGCCGACGC 2141
Oy 865 eGly----- 866
Db 2142 AGGCAAGCTGTGGGACGCGACACTATACCGCCCGCAAAACGTAACACAAATCGCT 2201
Oy 867 -----AsnAlaTyrAspIleGlyPheThrAsnAspThrLysIleGlyValTy 883
Db 2202 TTACTCGGAAACCGGCATTAATCCACCTTACCAACGAATTCGCTATTCGCGCGG 2261
Oy 883 IGIValArgTyrPalaSerProValIleGlnValArgValAspValAlaThrGlyValTy 903
Db 2262 CCGGTTACTGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2321
Oy 903 sGlu-----GluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
Db 2322 GAAAAACCGGAAGACGAATCCAAACGTTCCAAATTCAGCTCGGACGACGCTTC 2376

RESULT 26
LOCUS AX202493 2379 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 6 from Patent WO0152885.
ACCESSION AX202493
VERSION AX202493.1 GI:15392216
KEYWORDS
SOURCE
ORGANISM Neisseria gonorrhoeae.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE 1 (bases 1 to 2379)
AUTHORS Plizze M.C., Rappuoli R.C. and Giuliani M.C.
TITLE Outer membrane vesicle (omv) vaccine comprising N. meningitidis
JOURNAL seto group b outer membrane proteins
Patent: WO 0152885-A 6 26-JUL-2001;
Chiron Spa (IT)
FEATURES
source 1..2379
Location/Qualifiers
/organism="Neisseria gonorrhoeae"
/db_xref="taxon:485"
BASE COUNT 636 a 755 c 596 g 392 t
ORIGIN
Alignment Scores: 8.84e-08 Length: 2379
Pred. No.: 265.50 Matches: 195

```

```

Percent Similarity: 34.28% Conservative: 120
Best Local Similarity: 21.22% Mismatches: 348
Query Match: 5.62% Indels: 257
Db: 6 Gaps: 39
US-09-914-168-2 (1-919) x AX202493 (1-2379)
Oy 116 ThrProLeuSerLeuGluGlnLeuPheAlaGlnIleuSerThrGlnMetGlyIleAsnPro 135
Db 46 TCGCTTTGGCATTTGGCGCATTCACCATCCAAAGACATCCGTGTGCAAGGCTTCAGCGCT 105
Oy 136 AsnAsp-----TyrIleProGluTyrGlnGlyGlnIleProAsnSer 149
Db 106 ACCGAGCCGAGACCCGATTCACATCACTGCGCGCTCAAAAGCGGAGACACTACAC-- 162
Oy 150 GluValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
Db 163 -----GACACACGCGGAGTGCCATCATCAAAAGCCTGTAC 198
Oy 170 AlaArgLeuPheAsnAspGlyValAlaHisLysValPro-----ArgLeuLysAlaLys 186
Db 199 GCCACCGGTTCTTTGACGACGATGACAGTGCAGAACCTGCGGACGGGCACTTCTGTACC 258
Oy 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
Db 259 GTTATCGAACGCGCCACCATCGGCTCGCTCAACATCACCGCGCCAAATGCTGCAAAC 318
Oy 207 GluProTyrAlaHisLysLysAlaAlaLeuGlnLysP-----IleThrGlnIleSerAla 224
Db 319 GACGCC-----ATCAAGAAACCTCGAATCGTTCGGCGTGGCGGACGTGCAATAC 369
Oy 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
Db 370 TTATACGCGGACACATCAACCAAGCGAGTGGCGCGCTGAAGAAAGATACCTGGCGGT 429
Oy 237 -----ThrAlaLeuValAlaAlaArgValAlaGlyTyrTyrAsp 249
Db 430 GGCAAACTCAATATCCAAATACGCGCCCAAGTAAACCAACTCCGCCCAACCGCGTGAC 489
Oy 250 IAspLeuSerIleIleArgAsnSerIleGlyLysValAspValIleIleHisLysLeu 269
Db 490 ATGACATCATCGATTTGACGAGGCGCAATCCGCCAAATATCACCGCATCGAATTTGAA-- 546
Oy 270 GlyLysProValTyrIleAspTyrArgAlaValIleGluValArgGlyGluGlyAlaAsp 289
Db 547 GGCACCAAGCTATTCGACCGGCAACGTATGCG-----CAG 585
Oy 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisHis 309
Db 586 ATGTGCGTGCACGAGCGGCGCATTTGGACATGGCTGCACGAAACGCGGTTGACCGC 645
Oy 310 GlyLysTyrGlnThrLysLysAsnLeuIleGluAsnLysSerAlaGlnHisGlyTyrPhe 329
Db 646 CACAATTCGCCCAAGACATGAAAGAAAGTAACGACTTCTTACCAAGAAACCGCTACTTC 705
Oy 330 AspGlyArgTyrPheLysAspArgSerValAspValIleLeuProAspAsnThrAlaAsp 349
Db 706 GATTTCCGATCTCTCGATACCGACATCCAAACCAAGACAAACCAAGCGGACCATC 765
Oy 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIle 369
Db 766 AAATATCACGCTCCACGAGCGGACGCTTCGCTGGGCGCAAGTG-----TCGATT 816
Oy 370 AspProLysThrAsnGlnIleuThrThrAspProAspLysLeuProValLysArgGluLeu 389
Db 817 GAAGGGACACCAACGAA-----GTCCCAAGCGCGAA 849
Oy 390 LeuGluGlnLeuLeuThrValAsnMetGlyLysAlaTyrAsnLeuGlnIleValAlaArg 409
Db 850 CTGGAAATCTGTGACATGAAGCCCGCAATGTGTACCAAGCCGACGATGACCGCC 909
Oy 410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhe 429

```

```

Db 910 GTTTGGGT----- 918
Qy 430 PROGLIARGLUCINILEGINASPSGLINVALSERPHEGLUGINSErSerSerArg 449
Db 919 -----GAGATTGGAACCCGATGGCTGGCAGCATACAGC----- 960
Qy 450 ThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeu 469
Db 961 -----GAAATCAGCGGTACAGCGCGCTGGCAACGCCGGAACCA 999
Qy 470 ThrAspGlyIleLeuMetLysIleSerPro-----IleGlu 481
Db 1000 ACCGTCGATTTCTCTCGATCGAACCAGCGCGGAAATCTACGTCAACGAATCCAC 1059
Qy 482 PheSerAlaSerIleuIleGlnAspGlyLeuAsnIleuValAlaAlaLysAlaArgHis 501
Db 1060 ATCACCAGCAACAAACAAACCCCGGAGAA-----GTCTGCGCCGCGAATGGCGCAA 1113
Qy 502 LeuTyfAspMetProAspAsp-----ArgValIleuAla 512
Db 1114 ATGGATTCGGCGCTTACGACACCTCCACAGCTGCACAGCCTCAAGAGCGCGTGCAGCTT 1173
Qy 513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532
Db 1174 TTGGCTACTCTCGACCAACGTACAG-----TTTGAT 1203
Qy 533 AlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
Db 1204 GCCGTCGCCGCTTCCGCTACGCCCGACAAA-----GTCGATTTG----- 1242
Qy 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyfGlnSerLysValPro 572
Db 1243 AACATGAGCCTGACCGACAGCTCCAC----- 1269
Qy 573 LeuTyfValAlaPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeu-----Gly 590
Db 1270 -----GGCTCGCTCGATTCGATGAGCGCGCG 1293
Qy 591 TrpGlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArg 610
Db 1294 TGGGTTCAGATACCGGC-----TTGGTCATGTCGCCGCGCGATGCGAGGACACCTG 1347
Qy 611 AspGlyTyfGlnAlaGlnValGluLeuArgLeuSerGluAspLysGlyValLysLeu 630
Db 1348 TTCGGTACGGGCAAGTGGCGCGCTGGCGCTCGCGAAGCAAA-----ACAGCCTC 1401
Qy 631 TyfAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyf 650
Db 1402 AACGCTCGCTGTCGTTTACCGACCGCTTACGCGACAGCGGGTCCAGCTGGGCTAC 1461
Qy 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670
Db 1462 -----GATATTACGGAAGAACCTTCGACCGCGCAACACATCGACAGCTCAACAA 1515
Qy 671 HisGluIleSerArgSerIleIleGlnAsnGlyGlyTyfPasnArgThrTyfSerLeuArg 690
Db 1516 TATTAACACACACCGCGC-----GGCGGCGCGCTAAGAGATGGG-TATCCCGCT 1562
Qy 691 TyfArgLeuAspLysLeuLysThrGlnAlaProProGlu-----ThrTr 705
Db 1563 TACCGAATACGACCGCGCTAATTTGGCGTGGCGGCAACACCTGCACCTCAACACCTA 1622
Qy 705 PGIAspLeuPro-----ValAspPheValAsn-----GlyLys---ProSerGI 719
Db 1623 CAACAAGACCAACCAACGCTATGCCGCTTATCAGAAATACGCGCAAAACCGACGCGCC 1662
Qy 719 nGluAlaLeuAlaGlyValAlaValHisLysThrVal-----Al 733
Db 1683 AGACGCGACGCTTCAAGAGCTCTGTACAAAGGACCGCTGGCGGCGCACACAGAC 1742
Qy 733 aspAsnLeuValAsnProMetArgGlyTyfArgGlnArgTyfSerLeuGluValGlySe 753
Db 1743 CGACAGCGCGCTCATGGCCGACGCGCGCTAC----- 1773

```

```

Qy 753 rSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTy 773
Db 1774 -----CTGACCGCGCTAAATGCGGAATTCCTCCCGCGGACCAACTGCANFA 1823
Qy 773 rSerPheGlyAspAsnAlaTyfGlySerAsnArgAlaHisGlnMetThrGly----- 790
Db 1824 C-----TACTCCGCCACCCACCAACCAACCACTGGTTCTTCC 1859
Qy 791 -----GlyIleGlnAlaGlyTyfTrI 797
Db 1860 CTTAAGCAAACTTCACGCTGATGCTCGCGCGCAAGTCCGATTCGGCGGCGCTAC-- 1917
Qy 797 eTrpSerAspAsnPheAsnHisValProTyfArgLeuArgPhePheAlaGlyAspAl 817
Db 1918 -----GGCAGAACCAAGAAATCCCTTTGAAAACCTTCAGCGCGCGCTGGG 1970
Qy 817 nSerIleArgGlyTyfAlaHisAspSerLeuSerProIleSerAspLysGlyTyfLeuTh 837
Db 1971 TTGGTGGCGCGGCTACGAAACGCGCAGCTCGCGCG-----AAAGTATAGACGA 2021
Qy 837 rGlyGlyGlnValLeuAlaValGly-----ThrAlaGlyTyfAsnTyfGluPhe 854
Db 2022 ATACGGCGAAAAATACAGTACGGCGCACAAACCAACGCTCCGCGCAGCTGCT 2081
Qy 854 t-----LysAspLeuArgLeuAlaValPheGlyAspAl 865
Db 2082 CTTCGCCGATGCCCGGTGGCAAGACGACGACCTCGCTGACGCTGTGTGGCGGACCG 2141
Qy 865 eGly----- 866
Db 2142 AGCAGCGCTGTGGGACGCGACACTATACCGCGCGCAACCGTACCAACATCGGT 2201
Qy 867 -----AsnAlaTyfAspLysGlyPheThrAsnAspThrLysIleGlyAlaGly 883
Db 2202 TTACTCGGAACCGCGCTAATATCCATTCACACCAATTCGCTATTCGCGCGCGG 2261
Qy 883 IGIValArgTrpAlaSerProValIGlyValArgValAspValAlaThrGlyValTy 903
Db 2262 CGCGGTACCTGCGCTCGCTTTGGTCCGATGAAATTCAGTACGCTACCGCTGAA 2321
Qy 903 sGlu-----GluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
Db 2322 GAAAAACCGGAGACGAAATCCAAACGCTTCCATTCACGCTGGCAGCAGCTTC 2376

RESULT 27
AE010493/c 10029 bp DNA linear BCT 01-JUL-2002
LOCUS 10029 bp DNA linear BCT 01-JUL-2002
DEFINITION Fusobacterium nucleatum subsp. nucleatum ATCC 25586, section 35 of 197 of the complete genome.
ACCESSION AE010493 AE009951
VERSION AE010493.1 GI:19713160
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL MEDLINE
PUBMED 21886394
REFERENCE 2 (bases 1 to 10029)
AUTHORS Kapatali,V., Anderson,I., Ivanova,N., Reznik,G., Los,T., Lykidis,A., Bhattacharya,A., Battman,A., Gardner,W., Grechkin,G., Zhu,L., Vasileva,O., Chu,L., Kogan,Y., Chaga,O., Goldsman,E., Bernal,A., Larsen,N., D'Souza,M., Walunas,T., Pusch,G., Haselkorn,R., Fongstein,M., Kyrpides,N. and Overbeek,R.
Genome sequence and analysis of the oral bacterium Fusobacterium nucleatum strain ATCC 25586
J. Bacteriol. 184 (7), 2005-2018 (2002)

```

TITLE Direct Submission
JOURNAL Submitted (13-FEB-2002) Integrated Genomics, 2201 W. Campbell Park
Drive, Chicago, IL 60612, USA
FEATURES
source
/organism="Fusobacterium nucleatum subsp. nucleatum ATCC
25566"
/strain="ATCC 25566"
/sub_species="nucleatum"
/db_xref="ATCC:25566"
/db_xref="taxon:190304"
69.656
/gene="FNI907"
69.656
/gene="FNI907"
/EC_number="2.1.1.45"
/codon_start=1
/transl_table=1
/product="Thymidylate synthase"
/protein_id="AAL94006.1"
/db_xref="GI:19713161"
/translation="MKTITVAKKECKIEFEERKSKFTGYVAPVSKPEAEYIKYIKN
LHSDATHNCSAYKINNNGLEFPKVDDEPGSTAGKPMGDIINMEVSNLYIATRIY
GCIKAGAGLVNNVAKTAKLAITAEIIDFDIKDILFEIPIERLGEVEKLLAEYAE
VIDKSFLEIKYKVKINKNFYDNLNENYFNLDI"
complement(736..1809)
/gene="FNI908"
complement(736..1809)
/EC_number="3.1.4.46"
/codon_start=1
/transl_table=1
/product="Glycerophosphoryl diester phosphodiesterase"
/protein_id="AAL94007.1"
/db_xref="GI:19713162"
/translation="MKLSKGLVLTSSALFAANGKIIIAHRCASGILPEHTLES
KALFAQADVLEEDLAMSCKGLIYIHFDGLDLTVAKKPNRRADGRYVIDPT
WSELQLENTFTKQKOTAYPNRPFLMSKDEKLHFEIEIETGIEKSTGKI
GIPEIKAPWHONGKDTAKATLEVKYGTKSDMYLQTFDYNELKRTGELMP
KMGMDKLVLIVNDMHEIEEKDKGRWVNDYDMKEGAMKEIAKYAGVGVGGY
MLIDDKSKYKGNVITPMYKDIATTKMELHPYVRKDALPEFTVDNMYALLNKA
ATGVTDFDGLGVPLENKNK"
complement(1852..2850)
/gene="FNI909"
complement(1852..2850)
/gene="FNI909"
/EC_number="2.3.1.-"
/codon_start=1
/transl_table=1
/product="UDP-3-O-(3-hydroxymyristoyl) glucosamine
N-acyltransferase"
/protein_id="AAL94008.1"
/db_xref="GI:19713163"
/translation="MEYVTDIITFLNAEYKGEVEVSKLSFPFHSDEKSLTFPADE
KFLNLSQTKAKVILPIDIDPLEEGGIVAVDSPIVIMPKLHFEFRLKIEEMR
EDSAKIGENVDAIVNYIGHVYIGNNVKIFPVNTIGEGIIIDGTVISNYSIREM
ETGNKCVIOGCAVYSGDFGVKNGNNTIDIGIYIYEDVEIGANTTIDRGAGID
TIIKKYIKINDVIAHNDIIGENCLIIISQVGIAGSTTIGNNTLAGOVGVGHLEIG
DNPIAGOSGPGVNEANKIILSCHPIVDHREDMKIVAAKKCPPELLKRVKALEEK"
complement(2875..3348)
/gene="FNI910"
complement(2875..3348)
/gene="FNI910"
/codon_start=1
/transl_table=1
/product="periplasmic protein"
/protein_id="AAL94009.1"
/db_xref="GI:19713164"
/translation="MKLLVASVLLATSAFAKGVYVDSKAFQFSESTKKAQOSLE
SOAKVENEARKOEVALOAKDKDLDAKKAFFEKRSODFOAFNLSSGDKIN
KEDMAKKIELEVYKAIKKYVADSKYDIFEAELKVGEDITARYLKEEALK"
complement(3389..5425)
/gene="FNI911"

CDS
complement(3389..5425)
/gene="FNI911"
/codon_start=1
/transl_table=1
/product="Outer membrane protein"
/protein_id="AAL94010.1"
/db_xref="GI:19713165"
/translation="MVNLPKISVEVNNNOVASLTKNTLKIKEGAKFSTALLADFN
ALKETGFEDVILQPYSDGVYIVDYVEKENVVDLKEKVAINTLRDPSIVL
SVKFGNKRKVTSELDITQLKAGETVSRSEDAQRLLATGKFSERDAQVANG
KMAISFEVNPVTKSVITTTGNTLPTSTINSELTGKGSQVYNNLEDEBDKILG
QAGGYTLVNTDMSDNGTLLHISVEGVIRIEYKKNVYKONKRRPNDVYKTD
YVTDREITQPKIENKVEYDAVDNLMKLGIFKVKYKATKONKRRPNDVYKTD
KTDREIQGVAVGETGFLGTLSDKSMRGNQOFGFEPKSKNNTGRLDILIDE
IKTDVSWGVAAYRTSYGDEDSILEHEDITGFPNTGKGLKRNFLISGTKEYIK
EKHEDEKLOANNKRWYKERNKREITGVDDKWLMSIPYISYDTRNNVLPSTGF
YKFOVEAGHAGGYKSGNFGNATLELPTVHGLKRNNTFAKVVGVATNNKTESOF
VWGGNSLGRYDGFPGFSOKLVATIERKTLDNDIGLVFADAGRAMKONGRDPST
RNSRFGHIGTTAGVGRVNTPIGLPKEFGWGPVGNKDDGKPFYNNQGSF"
complement(5558..9085)
/gene="FNI912"
complement(5558..9085)
/EC_number="Hypothetical protein"
/codon_start=1
/transl_table=1
/product="Hypothetical protein"
/protein_id="AAL94011.1"
/db_xref="GI:19713166"
/translation="MDRIDESILNKIAPIREKKDLKLNINIEDIKAVIYSDPRLSI
KTKPKNSSEKGIELDNFLYISSKAGKNLSARLITIKGIPENIALSVNOKNT
DIIALASPKEKIDNIPDINIKAVENOKDIIKANIDSNVDPNVYNEKKLTKLYGN
KFTINLVNDKRIITNEGRIPETIYHTAVYLPVAKNNKTOIKELDKAKNSYFT
AKGANLVNDGEFTIDYEGKASISRKVKNENDILSDGKLENNKNNLSQOINDL
SLEYIKIERINGTVNLRKYGKNEANVITKIDISGIDYKRNPNLNVSENOVKI
KDSNNLISLKGVDYKNOKVANNISVNTKNDVAFKVEVEVNEANVEGOKIL
OGVAEDGSTITLTPSKDPYIKTGKSTKINSYVINDVNLIGKNNLEAKNDKIL
VLSAEKHEKVEYKADPGYILIGQIDVGVAKIAVAKGATNEKKLPDLADIER
NAENYADGASIASIDGDIIDRQYDGLGVKGVNLEKTLIDKNNKNTDLAKONLIS
NIDYVKNLVGKYIDIRKNTYNNLVKSNKIDVSKFSELPYTGIGINAKGNILNL
INENYERGINLENISLESTAKLILSNESGIPNGERIDVGLRSLNSNPLVVG
FVDLANISKLEDEDIRSLPYKLFHKMFVYFPEITIKSGSTELTAIEEYVGNLI
IDDAIVYDIPNNYRDFEFLIREOLKRPRTYVSTKIDDKOSKDKERKEVEMMLNK
LMPIDFVYTEREPILIDMDNENIYVPRYGCGLITDLNMLKKGKYYITGETLKEGYF
FYGTNFEQVDRALAVENNPENLNPVIFESITLNDDEYHNTKAGNOVREYS
SKTAVVGGDSALIVPNADNHIYSYDGNIEITFMKNLIIAGQAGGVGSTRYIK
RKFDLTKPVYIPEVKIINSDSVNRLLGQTTDNGFPEIYNNIKLEAKDNTRYIK
FKKANVRIITGIDVINKQTKMVDKSYREYVGVLEYVDDAKTIEIGVIVPKYRTD
PKKDIKRNHYHIGFRKRYRDFSEISF"

BASE COUNT 3102 a 1691 c 1027 g 4209 t

ORIGIN

Alignment Scores:

Score:	9.12e-07	Length:	10029
Percent Similarity:	262.50	Matches:	162
Best Local Similarity:	38.08%	Conservative:	111
Query Match:	22.59%	Mismatches:	265
DB:	5.35%	Indels:	179
	1	Gaps:	36

US-09-914-168-2 (1-919) x AE010493 (1-10029)

QY 311 LysTYRGluThrLysLysAsnLeuIleGlu---AsnAlaSerAlaGluHisGlyTyrPhe 329
|||||: |||||: ||| |||||
Db 5329 AAGTTTTCACACAGAGAGCTTATTAGCTTAAAGCCTTAAAGAAACACAGTTATTTT 5270
|||: |||: ||| |||||
QY 330 AspGlyVArgTirPleuAspArgSerValaAspValIleLeuProAspAsnThAlaAspVal 349
::: |||||: |||||
Db 5269 GAA-----GATGTAATTTCTTCAA----- 5252
|||: |||||
QY 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAsp-----GluValVal 365
::: ||||| ||| |||||
Db 5251 CCAAGTTTCTTATGATGCTGAGTGAAGATAGTGTGTTGTAAGAAAGAAATGTT 5192

QY 366 PhePheThrIleAspProLys-----ThrsnglnLeuThrAspProAspLys 382
Db 5191 GTGATTTATTGAAAGAAAAAGGTGTCATTAACACTTTTAAGAGAAATACAGATATAA 5132
QY 383 LeuProVal-----LysArgGluLeuLeuGlnLeu 393
Db 5131 TCATGTGATCTCATCATGTAATAATTTACTGGAATAAAGAGTTACTACATCAGAACTT 5072
QY 394 LeuThrValAsnMet-----GlyGlnAlaTyrAsnLeuGlnAlaValArgAlaLeu 410
Db 5071 TTGAGATTTTACACATTAAAGCAGGGAGATTTTTCACAGAGTTCAGGTTGAAGATCCT 5012
QY 411 SerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhePro 430
Db 5011 CAAAGAGATTTATAGCAGCTCGAAGAAATTTTCAGAGATTTAACACAGAT----- 4964
QY 431 GluArgGluGlnIleGlnAsnAspGln-----ValSerPheGluGlnSerSerSerSer 448
Db 4963 -----GCACAGATGACGAAATGCAAAATGCGCTTATCATCTTGA----- 4925
QY 449 ArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGlnThrValGlu 468
Db 4924 -----GTAGTTGAAAT-----CCAAATGATAAAAGTGAATA 4892
QY 469 LeuThrAspGlyIleLeuMetAspIleSerProIle-----GluPheSerAla----- 484
Db 4891 ATCAGCAGAAATATATACTATATCCAAAGATCTATTATGTCAGAAATGACACTTAACCT 4832
QY 485 -----SerAsnLeuIleGlnAspLys-----LeuAsnLeuVal 495
Db 4831 GGTTCAGCTCCAAATTTATACAAATCTTAAGAGATGAGATGATAAATTTAGGATTTGAT 4772
QY 496 AlaAlaLysAlaArgHisLeuTyrAspMetProAsp-----AspArgVal 510
Db 4771 CAAGCCCAAGATATCTTACTTTAGTAATATATACACATATGTACTGATGAATAATGGAAT 4712
QY 511 LeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArg----- 526
Db 4711 TTACATATTTCAATAGTTGAAGATATTGTAGAAGAAATGGAAGTTAAATAATGGTTACA 4652
QY 527 -----IleSerAspAlaValSerAlaValAlaArgAlaIle 538
Db 4651 AAACAAAAAGTAATAGAGAACAACCTAATGATGATGTTTGAATAAAGTAAACATATAT 4592
QY 539 LeuProAspGluSerGlu-----AsnGluValIleAspLeuProGluArgThrAla 555
Db 4591 ATGACACAGACAAATAGAAATACAACTCGAATAATATTATGTTAAGAAATATGATGCA 4532
QY 556 LeuAlaAsnArgLysThrProAlaAspVal-----TyrGlnSerLysLys 570
Db 4531 ACAGTTGATTAATTAATGATGATGGAATATTTCAAAAATGTTAATATGAGCAAGATCA 4472
QY 571 ValPro-----LeuTyrValPheValAlaSerAspLysProArg 583
Db 4471 ATTTCAGAGATCCAGAGAAATGATGATTTGATCTTTAATAGATGAGAGAAAGACTCT 4412
QY 584 AspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuValThrLysPhe 603
Db 4411 GAATTACAGAGTGGGTTGCTTATGTTCTGAACAGCA-----TTT 4370
QY 604 GluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlu-----Leu 619
Db 4369 TTAGGTACTTATCATTAAGATAGTAATCGAGAGGTAAATAATCAACAATTTGGTTTC 4310
QY 620 ArgLeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisPro 639
Db 4309 ACATTTGAAAAATCAATTAATAATTTACACAGTTTTCATTAAGATTTCTATGATCTTGG 4250
QY 640 LeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValAlaPheGlyHisSerThr 659
Db 4249 ATTAAGATACAGATAGAGTATCTTGGGATGG-----GGAGCTTACAGA 4205

QY 660 AsnGlyPheAspLeuSerThrArgThrLeuGlnHisGluIleSer-----Arg 675
Db 4204 ACTAGTTATGGATGAGATAGCATACTACTATCCATGAATATACATATAGCTTTCACA 4145
QY 676 SerIleIleGlnAsnGly-----GlyTyrAsnArgThrTyrSerLeuArgTyrArgLeuAsp 694
Db 4144 ACTAATATAGTAAGAGACTGTTGTAAGACTTTTACACTTATGTTAGAACAAAGTTGAA 4085
QY 695 LysLeuLysThrGln----- 699
Db 4084 TACATTAAGAAAAACATGAGATGAGAAATTAAGACAACAATTAATGTAATGTAAT 4025
QY 700 ---AlaProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysProSer 718
Db 4024 TATAAGAAAAAATAAATAGAGAGAA-----ATGAAAGGTGTAGATACAAATCTCG 3971
QY 719 GlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsn 738
Db 3970 TTGTGAGATATCTATCTTATCTATCTATCTATCTATCTATCTATCTATCTATCTAT 3917
QY 739 PrometArgGlyTyrArgGlnArgTyrSerLeuGlnValGlySerSerGlyLeuValSer 758
Db 3916 CCTACATCTGATCTGATCTATGGAATAATTTCAAGTTGAAGCTGACATGCTGTGATTAAG 3857
QY 759 AspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSer-----PheGly 776
Db 3856 TCAGAGAACTTTGGAATGCTACTTTGGAATTTGAGAACATATCATATTAAGATTTTAA 3797
QY 777 AspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyr 796
Db 3796 AATAATATATTT-----GCTTACAAAGTTGTGTGAGAGATCTACT----- 3755
QY 797 IleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyLysAsp 816
Db 3754 -----AACATACAAAGAAAGCAAAATTTCTGGTGGTGGAGCA 3713
QY 817 GlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeu 836
Db 3712 AATTCACATGAGAGATAT-----CATGAGAGATCTTT 3680
QY 837 ThrGlyGlnGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAsp 856
Db 3679 AAAGGAAGCAAAACTT-----GTGGCACTATTTGAATAAGAACACACTTAATGATATT 3623
QY 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr-----AspLys 871
Db 3622 ATGAGACTTGTGTATTTCTGATGAGCTAGACATGCAACAAATATGGAAGATCTCT 3563
QY 872 GlyPheThrAsnAspThr-----LysIleGly-----AlaGlyValGly 884
Db 3562 AGTTACACAGAGATAATAGTGTGTTTGACATTAACATAGAGCACTGCTGGGTGGGA 3503
QY 885 ValArgTTPAlaSerProValGlyGlnValArgValAspValAla-----ThrGlyVal 902
Db 3502 ATCAGACTTAATATCTCAATGACCATTTAAGATTTGATTTGTTGGCCAGTAGGAAT 3443
QY 903 LysGlnGluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
Db 3442 AAAATGATGATGATGAGATGAATTAATTTCTTCAATATATGGACATCAATTC 3392

RESULT 28
LOCUS NG081959 2379 bp DNA linear BCT 24-JAN-2002
DEFINITION Neisseria gonorrhoeae outer membrane protein (omp85) gene, complete cds.
ACCESSION U081959
VERSION U081959.1 GI:1766041
KEYWORDS
SOURCE
ORGANISM Neisseria gonorrhoeae.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
REFERENCE 1 (bases 1 to 2379)

AUTHORS Manning, D.S., Reschke, D.K. and Judd, R.C.
 TITLE Omp85 proteins of *Neisseria gonorrhoeae* and *Neisseria meningitidis* are similar to *Haemophilus influenzae* D-15-79 and *Pasteurella multocida* Oma87
 JOURNAL Microb. Pathog. 25 (1), 11-21 (1998)
 MEDLINE 98379445
 PUBMED 9705245
 REFERENCE 2 (bases 1 to 2379)
 AUTHORS Reschke, D.K., Manning, D.S. and Judd, R.C.
 TITLE Direct Submission
 JOURNAL Submitted (11-DEC-1996) Division of Biological Sciences, University of Montana, Health Sciences 104, Missoula, MT 59812-1002, USA
 FEATURES
 Location/Qualifiers
 1..2379
 /organism="Neisseria gonorrhoeae"
 /strain="FA19"
 /db_xref="taxon:485"
 1..2379
 /gene="omp85"
 1..2379
 /note="omp85"
 /codon_start=1
 /transl_table=11
 /product="outer membrane protein"
 /protein_id="AAC17600.1"
 /db_xref="GI:1766042"
 /translation="MKIKQTASALMMGLISPLAFADFTIIDIIEVGLIORTPESTVENYLPVAVDVTYNDTHGSAIKSLVATGFEEDVRYETAGLLIYVCTPTGSLMTAKLMONDAIKKMLSEFGLAQSOYFNQATLNOVAGLKEEYSGRLINQIDTPKVKLANRVIDITIDEGKSAKITDIEFEENQOYSDKLMKROSKITGTLTSLTSDRDRKFAADMEKVTDFYONNGYFDFRILDTDIOTNDKRTQKITVHEGGRFMGKVSIEDDTNEVPAKLEKLLTMKPKKTERQOMTAIVCEIONRMSAGYATSEISVPLPNAQKTVDPVLIHERGRKTYNEHITGNNKTRDEVRELRRLQMSAPYDTSKSRERELLGYRDNQFDPAVPLAGTPDKVYDLMNLSLESTGSLDSAGVODTGLVMSAGVQDWLFGRKSAALASRSKRTLPNGSLSTDEYFTADGVSGLDIYGRKAFDRKASTSVKQYKTAAAGGVNMGIPVTEYDRYVNFGLAEHLVNTYKNAKPRADFIKQYGTDDAGDSERGLYKGTGVGMGRKTDALMPTRKGYLTGVNAETALPGSKLOYSAATHNDPFRPLSKFTLMLGEVGIAGYGRKEIPFEENFGCGSGVNGYSGTIGKPVYDEYEGKLSYSGGNKANVASAILLPMPGAKDARTVRLSLRADSGSWDGRVTYTAENGNNKSVSEMAHKSTFTNELKYSAGCAVTMLSPGPMKFIYAPLKKRPEDEIORFOLFCTTP"

BASE COUNT 636 a 752 c 592 g 399 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2, 11e-07 Length: 2379
 Score: 259.50 Matches: 192
 Percent Similarity: 33.958 Conservative: 120
 Best Local Similarity: 20.898 Mismatches: 351
 Query Match: 5.49% Indels: 257
 DB: 1 Gaps: 37

US-09-914-168-2 (1-919) x NC081959 (1-2379)

OY 116 ThrProLeuSerLeuGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
 DB 46 TCCCTCTTGGCATTTGCCACTTCCACATTCACCAAGACATCCGTGTCGAAGCTTGCGGCT 105
 OY 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149
 DB 106 ACCGACCCAGACACCGCTATTCACTACCTGCCCCGTCAAAAGTGGGACACCTTCAACAC--- 162
 OY 150 GluAlaValValProProThrLeuGluProGluLysProGlyIleuIleLysArgLeuTyr 169
 DB 163 -----GACACACACGGCAGTCCCATCATCATCAAAAGCCTGTAC 198
 OY 170 AlaArgLeuPheAsnAspGlyVal-----ArgValLeuAla 512
 DB 199 GCCACCGCTTTCTTTGACGACGCTACGACGTGAACCTGGCAGCTCTTCTGCTGACC 258
 OY 178 -----AsnLysValProArgLeuLysAlaLysPheTyrGlnSer 190
 DB 259 GTTATCGTATGCCCTACCATCGGCTGCTCAACATCACCGGCGCAAAATGCTGTCAGAAC 318

OY 191 SerGln-----SerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyr 209
 DB 319 GACGCCATCAAGAAACCCTGCATATCGTTGGCGCTGGGCACTGCCAATTAATTAACTAG 378
 OY 210 AlaAsnIleLysAlaIleLeuGluAspIleThrGlnGluSerAlaMetAspLeuAsnGly 229
 DB 379 GCGACACTCAACCAAGCAGTGGCGCGCTCGAAGAGAAATAT-----CTGGGGGCGC 429
 OY 230 SerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
 DB 430 GCGAATCATATATCAAAATACACGCCCAAAATGTAACCAAACTGCCGCAACCGCGCTGAC 489
 OY 250 IleAspLeuSerIleIleArgAsnSerIleGlyValAlaValAlaValIleIleHisAspLeu 269
 DB 490 ATTCACATCACTATTCGACGAGCAAAATCCGCCAAATATCACCGACATTCGAATTGAA--- 546
 OY 270 GlyGluProValTyrIleAspTyrArgAlaValAlaValArgIleGlyGluAlaAspAsp 289
 DB 547 GGCACCAACAGTCTATTCGACGCGCAAACTGATGCGG-----CAG 585
 OY 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisHis 309
 DB 586 ATGTCGCTGACCCGAGCGGCGATTTGGACATGCTGTACAGAGCGACCGGTTGACCGC 645
 OY 310 GlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGlnHisGlyTyrPhe 329
 DB 646 CAGAAATTCGCCCAAGACATGAAAGAAAGTAACCGACTTTCACCAAGAACACGCTACTTC 705
 OY 330 AspGlyArgTyrPheAspArgSerValAspValIleLeuProAspAsnThrAlaAspVal 349
 DB 706 GATTCGCTGCTCCGATACCGATCCCAATCAACACCAAGACAAACACGACGACCATC 765
 OY 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIle 369
 DB 766 AAATATCCGCTCCAGAAAGCGGACGTTCCGCTGGGCAAAAGT-----TTCGATT 816
 OY 370 AspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeu 389
 DB 817 GAAGGGGACACCAACGAA-----GTCCCAAGGGCGGAA 849
 OY 390 LeuGlnGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValAlaGla 409
 DB 850 CTGGAATAACTGCTGACCATTAAGCCCGCAATGTGTACAGACCGACGACGATGACCGGC 909
 OY 410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhe 429
 DB 910 GTTTTGGCT----- 918
 OY 430 ProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArg 449
 DB 919 -----GAGATTCAGAACCGCATGGCGCTGGCAGGTACGATACAGC----- 960
 OY 450 ThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeu 469
 DB 961 -----GAAATCAAGCTTACAGCGCGCTCGAAGCGCGGAGCAAA 999
 OY 470 ThrAspGlyIleLeuMetAspIleSerPro-----IleGlu 481
 DB 1000 ACCGTGCAATTCGCTGCTGACATTCGACACCGGCGAGAAAATCTACGTACAGAAATCCAC 1059
 OY 482 PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaValAlaArgHis 501
 DB 1060 ATCACCGGCACACACAAACCGCGACGAA-----GTGCGGCGCGCGGAATTCGGCCAA 1113
 OY 502 LeuTyrAspMetProAspAsp-----ArgValLeuAla 512
 DB 1114 ATGGAATCCGGCTTACGACACCTCCAAAGTGCACACCTCCAAAGACCGCGTGCAGCTT 1173
 OY 513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532
 DB 1174 TTGGGCTACTTGCACAAAGCTACAG-----TTTGAT 1203

```

OY 533 AlaValAlaArgAlaIleuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
    ||||| ||| |||||
Db 1204 GCCGTCGCCGCTGCGGTACGCCGCAACA-----GTCGATTTG----- 1242
OY 553 ArgThrAlaLeuAlaAsnArgIleThrProAlaAspValIleGlnSerLysLysValPro 572
    :|||: ||| |||
Db 1243 AACATGAGCCTGACCGCAAGCTTCCACC----- 1269
OY 573 LeuTyValIlePheValAlaSerAspLysProArgAspGlyGlnIleGlyLeu-----Gly 590
    ||| ||| |||
Db 1270 -----GCTCGCTCGACTTGAGCGCGG 1293
OY 591 TrpGlySerAspThrGlyThrArgLeuValThrLysPheGlnIleHisAsnLeuIleAsnArg 610
    ||| ||||| |||||
Db 1294 TGGGTTACAGATACGAGCG-----TTGTCATGTCGCGCGCGCATGCGAGCAACACTG 1347
OY 611 AspGlyTyTyrGlnAlaGlyValGlnLeuArgLeuSerGlnAspLysLysGlyValLysLeu 630
    ||| ||| ||||| |||
Db 1348 TTCGGTACGGGCAAGTGGCGCGCGCTCGCGCAAGCAAA-----ACCACGCTC 1401
OY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
    :|||: ||| |||
Db 1402 AACGGTCGCTGCTGCTTACCGACCGCTACTTACGCGACAGCGGGTACAGCTGGGCTAC 1461
OY 651 GlnGlnGlnValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670
    :|||: ||| |||
Db 1462 -----GATATTACGGAAAGCCTTCGACCGCGCAACATCATGACGCTCAACAA 1515
OY 671 HisGlnIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThrTyrSerLeuArg 690
    :|||: ||| |||
Db 1516 TATTAACCAACACACCGCC-----GCGCGCGCGCTAAGATGG-TATCCCGCT 1562
OY 691 TyrArgLeu-AspLysLeuLysThrGlnAlaProProGlu-----ThrTyr 705
    |||||: ||| |||||
Db 1563 TACCGAATACAGACCGGCTCATTCGGCGTGGCGCGGCAACACTGACCTCAACACTTA 1622
OY 705 PclnAspLeuPro-----ValAspPheValAsn-----GlyLys--ProSerG 719
    :|||: ||| |||||
Db 1623 CAACAACACCAACCAACGCTATGCGCATTCGCAACATCATGACGCAAAACCGACGCGC 1682
OY 719 nGlnAlaLeuAlaGlyValAlaValHisLysThrVal-----Al 733
    :|||: ||| |||||
Db 1683 ACACGGCAGCTTCAAAGCGCTGCTGACAAAGCGACTGCTGGGTGGGGCGCAACAAGAC 1742
OY 733 AspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySe 753
    :|||: ||| |||||
Db 1743 CGACAGCGGCTTATGCGCGACGCGCGGCTAC----- 1773
OY 753 tSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyLysSerGlyValTyr 773
    :|||: ||| |||||
Db 1774 -----CTGACCGCGGCTAATCCCGAATCGCCCTGCCGCGACGCAAACTGCATTA 1823
OY 773 tSerPheGlyAspAsnAlaIleTyrGlySerAsnArgAlaHisGlnMetThrGly----- 790
    :|||: ||| |||||
Db 1824 C-----TACTCGCCACCCACACCAACCAACCTGGTCTTCCC 1859
OY 791 -----GlyIleGlnAlaGlyTyrIle 797
    :|||: ||| |||||
Db 1860 CTTAAGCAAAACCTTCACGCTGATGCTCGCGCGCGCAAGTCGCGATTCGCGGCGGCTAC-- 1917
OY 797 eTrpSerAspAsnPheAsnHisValProTyArgLeuArgPhePheAlaGlyLysAspG 817
    :|||: ||| |||||
Db 1918 -----GGCAGAACCAAGAAATCCCTTCTTGAATACTTACGCGCGCGCTGGG 1970
OY 817 nSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyLysLeuThr 837
    :|||: ||| |||||
Db 1971 TTCCGGTGGCGGTACGAAGCGCGACGCTCGGCGC-----AAATGTGATGACGA 2021
OY 837 tGlyGlyGlnAlaLeuAlaValGly-----ThrAlaGlnTyTyrAsnTyrGluPhe 854
    :|||: ||| |||||
Db 2022 ATACGGCGCAAAATCATGCTACGCGCGCAACAAACCAACGCAACGCTCGCGCAAGCTGCT 2081
OY 854 t-----LysAspLeuArgLeuAlaValPheGlyAspIle 865

```

```

Db 2082 CTTCCCGATGCGCGGTGGGAAAGACGCAACGCTCGCGCTGACCTGTTTGGGACGC 2141
OY 865 eGly----- 866
Db 2142 AGCAGCGGTGGGACGCGCAACACTTATCCGCGCGCAAAACGTAACAACTACGCT 2201
OY 867 -----AsnAlaTyAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyVa 883
    :|||: ||| |||||
Db 2202 TTACTCGAAACCGCGCATTAATCACCCTTACCAACGAATTCGCTATTCGCGCGCGG 2261
OY 883 tGlyValArgThrPheAlaSerProValGlyGlnValArgValAspValAlaThrGlyVal 903
    :|||: ||| |||||
Db 2262 CGCGGTACCTGCGCTCGCGCTTGGCGCGCATGAATTCATCTACGCTACCGCGCTGA 2321
OY 903 sGlu-----GluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
    :|||: ||| |||||
Db 2322 GAAAAAACCGGAAAGCAATTCACAGCTTCCAAATTCAGCTCGGACGAGCTTC 2376

RESULT 29
AX155439
LOCUS AX155439 2394 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 2 from Patent WO0138350.
ACCESSION AX155439
VERSION AX155439.1 GI:14536770
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE
1 (bases 1 to 2394)
AUTHORS
Giuliani,M.M., Pizza,M., Rappuoli,R. and Holst,J.
TITLE
85kDa neisserial antigen
JOURNAL
Patent: WO 0138350-A 2 31-MAY-2001;
Chiron Spa (IT) ; Statens Institutt for Folkehelse (NO)
FEATURES
source
1..2394
/organism="Neisseria meningitidis"
/db_xref="taxon:487"
BASE COUNT 657 a 743 c 583 g 411 t
ORIGIN

Alignment Scores:
Pred. No.: 2.29e-07 Length: 2394
Score: 259.00 Matches: 191
Percent Similarity: 33.48% Conservative: 121
Best Local Similarity: 20.49% Mismatches: 342
Query Match: 5.48% Gaps: 278
DB: 6
US-09-914-168-2 (1-919) x AX155439 (1-2394)
OY 116 ThrProLeuSerLeuGluGlnLeuPheAlaGlnGlnSerThrGlnMetGlyIleAsnPro 135
    :|||: ||| |||||
Db 46 TCGCCTTTGGCACTTGGCGACTTCACCATCCAAACATCCCGCTGCAAGCTTGACAGCT 105
OY 136 AsnAsp-----TyrIleProGluTyrGlnGlyGlnGlnProAsnSer 149
    :|||: ||| |||||
Db 106 ACCGAGCGGATACCGTATTCACCTACCTGCGCGTCAAGTGGGAGACCTACAC-- 162
OY 150 GlnValValAlaProProThrIleGlnGluProGluLysProGlyLeuIleLysArgLeuTyr 169
    :|||: ||| |||||
Db 163 -----GACACACAGCGGAGTGCATCATCAAAAGCTGTAC 198
OY 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
    :|||: ||| |||||
Db 199 GCCACCGGTTTCTTGGACGAGCTACGCGTCAAACTGGCGGAGCGGAGCTCTGTCGACC 258
OY 187 PheTyGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
    :|||: ||| |||||
Db 259 GTTATGCAACGCCACCATCGGCTCGCTCAACATCACGCGGCAAAATGCTGCAAAAC 318
OY 207 GluProTyAlaAsnIleLysAlaIleAlaLeuGluAsp-----IleThrGlnGlnSerAla 224

```

```

Db 319 GACGCC-----ATTAAGAAAAAACCCTCGAATCCTTGGGGCTGGCCGACCTCGCAATAC 369
Oy Metasp-----LeuAnGlySerIleProAlaArgIn----- 236
Db 370 TTTAATCAGGGGACACTAATCAGGAGCTCCGGCCCTGAAAGAGAATACCTCGGGCC 429
Oy 237 -----ThrAlaLeuValAlaAlaArgAlaValAlaGlyTyrAsp 249
Db 430 CGCAAACTCAATATCGAATACCAATACCAATACCAATACCAATACCAATACCAATACCAATAC 489
Oy 250 IleAspLeuSerIleIleArgAsnSerIleGlyValAlaAspValIleIleHisAspLeu 269
Db 490 ATCGCATCAGCATTCAGGACGAGGCAATCCGCCAATAATCACCAGCATCGAATTGAA--- 546
Oy 270 GlyClnProValTyrIleAspTyrArgAlaValAlaGlyValAlaGlyValAlaAspAsp 289
Db 547 GGCACCAAGCTATTCGACCGCAACTGATCCG----- 582
Oy 290 LysAlaPheThrThrValAlaAspClnValProLeuLeuIleGlyAspVal----- 306
Db 583 -----CAATGTCTCCCTGACCGCAAGCGCGCATTTGGACATGG 618
Oy 307 -----PheHisIleGlyLysTyrGlyThrLysLysAsnLeuIleGlu 320
Db 619 CTGACACGACCAACCAATTCACACGAGAAATTTGCCCAAGATATGAAAAAGTAACC 678
Oy 321 AsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgTyrLeuAspArgSerValAspVal 340
Db 679 GACTTCTACCAAAATTAAGGGCTACTTCGATTCGATCTCGATACCGACATCCAAACC 738
Oy 341 IleLeuProAspAsnThrAlaAspValSerIleTyrAspThrGlyThrGlnTyrArg 360
Db 739 AACGAAGCAAAACCAACCAAGACCATCAAAATTCACCGTCACGAAGCGGACGTCCTCGT 798
Oy 361 PheAspGlnValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
Db 799 TGGGGCAAGTC-----TTCATCGAAGCGCACACCAACGA----- 834
Oy 381 AspLysLeuProValLysArgGlnLeuGlnLeuGlnLeuThrValAsnMetGlyGlu 400
Db 835 -----GTCCCAAGCCGAACTGGAAAAAAGTGTGACCATGAAACCGCGCAAA 882
Oy 401 AlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrHe 420
Db 883 TCGTACGAAACCCACGACATACGCGCTTTGGGT----- 918
Oy 421 AsnMetValAsnThrGlnIleValPheProGlnArgGlnIleGlnAsnAspGlnVal 440
Db 919 -----GACATTCAACACCGCATGGGC 939
Oy 441 SerPheGlnGlnSerSerSerArgThrClnProAlaGlnValAspClnSerThrLeu 460
Db 940 TCGCGAGGCTACCGCATACGC-----GAATACAGCGTA 972
Oy 461 GlnProValIleGlnThrValGlnLeuThrAspGlyIleLeuMetAspIleSerPro--- 479
Db 973 CAGCGGCTCCGGAACCGTAACCAAAACCGTCGATTCGTCACATCGAAGCGGGC 1032
Oy 480 -----IleGlnPheSerAlaSerAsnLeuIleGlnAspLysLeu 492
Db 1033 CGCAAAATCTACGTCACGAATATACATACACCGGCAACCAACCAACCGGACAGCA--- 1089
Oy 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp----- 508
Db 1090 ---GTCTGCTCCCGCTGAATTCAGCCAAATCGCAATCCGACCTTACAGACACTCCAAAGCTG 1146
Oy 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAlaAsnArgSerIle 523
Db 1147 CAACGTTCCAAAGAGACCGCTGAGCTTTTGGGTACTTGGCAATGTCCAG----- 1197
Oy 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspClnSer 543

```

```

Db 1198 -----TTTGATGCTGTCCCGCTTGGCGGACCGCCGCAAA--- 1233
Oy 544 GlnAsnGlnValIleAspLeuProGlnArgThrAlaLeuAlaAsnArgLysThrProAla 563
Db 1234 ---GTGATTTG-----AACATGAGTGTGACCGAAGCTTCCACC----- 1269
Oy 564 AspValTyrGlnSerLysValProLeuTyrValPheValAlaSerAspLysProArg 583
Db 1269 ----- 1269
Oy 584 AspGlnGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr 601
Db 1270 ---GTTTCCCTGGATTTGAGCGCGGTTGGTTCAAAATACCGGG-----TGGTCATG 1320
Oy 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlnLeuArgLeu 621
Db 1321 TCCGACAGCGCTTCCCAAGAACACCTGTCGTCACGGCAAGTCGGCGACATCGCGCC 1380
Oy 622 SerGlnAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
Db 1381 TCCAGGAGCAAA-----ACCACGCTTAACGGCTCGCTGTCTACTGACCGCTACTTC 1434
Oy 642 AspClnLeuArgAlaThrLeuGlyTyrGlnGlnGlnValPheGlnHisSerThrAsnGly 661
Db 1435 ACGGCACAGCGGCTCAGCCTGGCTAC-----GATGTTACGAAAGCCTTCGACCCG 1488
Oy 662 PheAspLeuSerThrArgThrLeuGlnHisGlnIleSerArgSerIleIleGlnAsnGly 681
Db 1489 CGCAACAGCATCGACACATCAACATATTAACCAACCAACCAACCAACCAACCAACCAAC 1536
Oy 682 GlyTyrAsnArg-----ThrTyrSerLeuArgTyrAlaGlyLeuAspLysLeu 696
Db 1537 GCAGGATTCGCGATGACGCTGCTTACCGAATACGACGCGCGTAATTCGCTTGGTG 1596
Oy 697 LysThrGlnAlaProProGlnThrThrGlnAspLeuPro-----ValAspPheVal 713
Db 1597 GCACAACACCTGACCGCTACACCTACACAAACCGCCCAACACTTGGCGACTTATC 1656
Oy 714 Asn-----GlyLys---ProSerGlnGlnAlaLeuLeuAlaGlyValAlaValHisLys 730
Db 1657 AAGAAATACGCGCAAAACCGACGCGACGCGAGCTTCAACGCTGCTGCTCAAAAGT 1716
Oy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
Db 1717 ACCCTGCGCTGGGGGCCCAACAAACCGACGCGCTTATGCGCGGCGGTAC--- 1773
Oy 745 GlnArgTyrSerLeuGlnValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
Db 1774 -----CTGACGGGCGTGAACCGCAATC 1797
Oy 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
Db 1798 GCCCTGCTGGCGACCAACTGCAATAC-----TACTCCGCC 1833
Oy 785 AlaHisGlnMetThrGly----- 790
Db 1834 ACCCAACCAACCAACCTGCTTCCCTGAGCAAAACCTTCACGCTGATGCTCGCGGC 1893
Oy 791 -----GlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArg 808
Db 1894 GAATCGGCAATCGCGGCGGTAC-----GGCAAGAACCAAGAAATCCCTTCTTT 1944
Oy 809 LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
Db 1945 GAAAACTTCTACGCGCGCGCTGGTGGTGGTGGCGGATACGAACGCGGACGCTCGGT 2004
Oy 829 ProIleSerAspLysGlyTyrLeuThrGlyGlnValLeuAlaValGly----- 845
Db 2005 CCG-----AAAGTCTATGACGAATACGCGCAAAATAATCACTACGCGCGCAACAAA 2055
Oy 846 ThrAlaGlnTyrAsnTyrGlnPheMet-----LysAsp 856
Db 2056 AAAGCAACGCTTCGCGGAGCTGCTTCCCGATGCCCGCGGCAAGACGCGGACAC 2115

```



```

OY 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543
DB 1198 -----TTGATGCCCTGCCGCTTCCGATGCCGCGGCGAAGACGGCGCAC 1233
OY 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
DB 1234 -----GTCGATTGG-----AACATGAGCCTACCCCAACCTTCCACC----- 1269
OY 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583
DB 1269 ----- 1269
OY 584 AspGlyGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr 601
DB 1270 ---GGCTCCCTGCACTTGACCGCGCGCTGCTACAGATACCGGC-----CTGGGCAAG 1320
OY 602 LysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyValGluLeuArgLeu 621
DB 1321 TCCGCGAGCGCTTCCCAACACACCTGTCGGTACGGGCAAGTGGCCCGCTGGCGCGC 1380
OY 622 SerGluAspLysArgLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
DB 1381 TCACGACAGCAAA-----ACCACGCTCAACGGCTCGCTTACCGACCGCTACTTC 1434
OY 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661
DB 1435 ACGGACAGACGGGCTGACCTGCGCTAC-----GATGTTACGGAAACCTTCGACCGC 1488
OY 662 PheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly 681
DB 1489 CGCAAGCATCGACGACGATCAACACATTAACACACACGCA-----GGC 1536
OY 682 GlyTyrPasnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696
DB 1537 GCAGGATCCGATGAGCTGCTGCTTACCGAATACGACGCGGTGAATTCGCTTGGTG 1596
OY 697 LysThrGlnAlaProProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713
DB 1597 GCAGAACACCTGACCGCTCAACACCTTCAACAACAAAGCCCAAACTAGTCCGACTTTATC 1656
OY 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730
DB 1657 AAGAAATACGGCAAAACGACGACGACGACGCTCAAGCGCTGATACAAAGT 1716
OY 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgLysTyrArg 744
DB 1717 ACCGTGCGCTGGGGCGCAACAAACGACGCGCTTATGCGCGCGCTAC--- 1773
OY 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
DB 1774 -----CTGACGGCGCTGAACGCCCGCAATC 1797
OY 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaIleTyrLysSerAsnArg 784
DB 1798 GCCCTGCCGCGCAACCACTGCATAC-----TACTCCGCC 1833
OY 785 AlaHisGlnMetThrGly----- 790
DB 1834 ACCCAACAAACCTGCTTCCCTTAAGCAAAACCTTACACGCTGATGCTGGCGGC 1893
OY 791 -----GlyIleGlnAlaGlyTyrIleTyrPheAspAsnPheAsnHisValProTyrArg 808
DB 1894 GAAGTCGCGATTGCGCGCGCTAC-----GCCAGAACCAAAACAAATTCCTTTT 1944
OY 809 LeuArgPhePheAlaGlyLysArgLysPheGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
DB 1945 GAAACTTCTACGCGCGCGCTGCGTGGTGGCGGAGTACGAAACCGCGACGCTCGGT 2004
OY 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly----- 845
DB 2005 CCG-----AAAGTGTATGACGAATACGCGAATAAATACGCTACGCGCGCAACAA 2055
OY 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856

```

```

DB 2056 AAGCCCAACGTCCTCCGCGAGCGCTCTCCGATGCCGCGGCGAAGACGGCGCAC 2115
OY 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
DB 2116 GTCCGCTGAGCTGTGTGGCGACGACGACGCTGTGGACGGCAAAACCTACGACGAC 2175
OY 870 ----- 870
DB 2176 AACACGATTCGCGGCGCGGCGGCAAGGTTCAAAACATTTACGGCGCGGCAATCCAT 2235
OY 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValArgThrAlaSer 889
DB 2236 AATCCACCTTTTACACGAATTTGCGCTATTCGCGCGCGGCTTACCTGGCTTCG 2295
OY 890 ProValIleGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
DB 2296 CCTTAGCGCCGATGAATTCAGCTACGCTTACCGCGCTGMAAAGAAACCGAAGACGAA 2355
OY 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
DB 2356 ATCCAACGCTTCCATTCCAACTCGGACAGAGTTC 2391
RESULT 31
AX202489 AX202489 2394 bp DNA linear PAT 30-AUG-2001
LOCUS Sequence 2 from Patent WO0152885.
DEFINITION AX202489
ACCESSION AX202489
VERSION AX202489.1 GI:15392215
KEYWORDS
SOURCE
ORGANISM
Neisseria meningitidis.
Neisseria meningitidis.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.
REFERENCE
1 (bases 1 to 2394)
Piazza, M.C., Rappunli, R.C. and Giuliani, M.C.
Outer membrane vesicle (omv) vaccine comprising N. meningitidis
serogroup b outer membrane proteins
Patent: WO 0152885-A 2 26-JUL-2001;
JOURNAL
Chiron Spa (IT)
FEATURES
Location/Qualifiers
source 1..2394
/organism="Neisseria meningitidis"
BASE COUNT 657 a 743 c 583 g 411 t
ORIGIN
Alignment Scores:
Pred. No.: 2 29e-07 Length: 2394
Score: 259.00 Matches: 191
Percent Similarity: 33.48% Conservative: 121
Best Local Similarity: 20.49% Mismatches: 342
Query Match: 5.48% Indels: 278
DB: 6 Gaps: 40
US-09-914-168-2 (1-919) x AX202489 (1-2394)
OY 116 ThrProLeuSerLeuGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
DB 46 TCGCCTTTGGCAGCTTGGCGACATCCACATCCAGACATCCGCGTGAAGGCTTGACGCT 105
OY 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149
DB 106 ACCGAGCGGAGTACCGTATTCACATACCTGCGCCGCAAGTCCGCGACACCTACAAAC--- 162
OY 150 GluValValValProProThrLeuGluProGluLysProGlyLysIleLysArgLeuTyr 169
DB 163 -----GACACACACGCGAGCTGCCATCATCAAAAGCTGTAC 198
OY 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
DB 199 GCGACCGGTTCTTGTGACGAGCTACGCGTCGAACACTCGGAGCGGACGCTCTGCTGACC 258

```

QY	187	PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnTyrThr	206
Db	239	GTTATTCGAAGCGCCACCACATCGCTCGCTCAACATTCACCGCGCAAAATATGCTGCAAAAC	318
QY	207	GluProTyrAlaAsnIleIleGlyAlaIleGluGlnuAsp-----IleThrGlnGlnSerAla	224
Db	319	GACGCGC-----ATTAAAGAAAACCTCGCAATCGTTTCGGGCTGGCGCGCATGCGCAATAC	369
QY	225	MetAsp-----LeuAsnGlySerIleProArgGluLeuArgGln-----	236
Db	370	TTTAAATCAGCGGACACATCAATCAGGACGATGCGCGGCTCGAAAGAAAGAAATCTCGGGCGG	429
QY	237	-----ThrAlaLeuValAlaIalaArgAlaValGlyTyrTyrAsp	249
Db	430	GCGCAAACTCAATATTCCAATATCCAGCCCGCAAACTCAACACTCGCGCCGCAACCGCTGCGAC	489
QY	250	IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu	269
Db	490	ATCGACATCAACGATTTGACGAGGCGGAATCCGCCCAAAATTCACCGACATCGAATTGAA---	546
QY	270	GlyGluProValTyrTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp	289
Db	547	GGCACACCAAGTCTATTCGACCGCAAACTGATGCGG-----	582
QY	290	LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal-----	306
Db	583	-----CAAAATGCCCTGACCGCAAGCGCGCATTTGACATGCG	618
QY	307	-----PheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGlu	320
Db	619	CTGACACCAACCAACCAATTCACAGCAGCAGAAATTTCCCAAGATATGAAAAAGTAACC	678
QY	321	AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPheAspArgSerValAspVal	340
Db	679	GACTTCCTACCAAAATTAACGGCTACTTGATTCGGTCTCGATCCAGCATCCCAACACC	738
QY	341	IleLeuProAspAsnThrAlaAspValSerLeuLeuIleTyrSphTrpGlyThrGlnTyrArg	360
Db	739	AACGAAACCAAAACCCAGACACCATCAAAATCCAGCTCAGCAAGCGGACCTTTTCCT	798
QY	361	PheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro	380
Db	799	TGGGGCAAAAGTC-----TTCATTCGAAAGCGCACACCAACGA-----	834
QY	381	AspLysLeuProValLysArgGluLeuLeuGlnGlnLeuLeuThrValAsnMetGlyGlu	400
Db	835	-----GTCCCAAAAGCGCAACTGGAAAAAAGCTGCTGACATGAAGCCCGGCMAA	882
QY	401	AlaTyrAsnLeuGlnAlaValArgAlaIleuSerAsnAspLeuIleAlaThrArgTyrPhe	420
Db	883	TGGTATCGAAGCCCGACAGATACCGCGGTTTTGGGT-----	918
QY	421	AsnMetValAsnThrGluIleValPheProGluArgGlnGlnIleGlnAsnAspGlnVal	440
Db	919	-----GAGATTCAGAACCGCATGGCC	939
QY	441	SerPheGlnGlnSerSerSerArgTyrGluProAlaGlnValAspGlnuSerThrLeu	460
Db	940	TTCGCGAGCTACGCATACACG-----GAAATCCACCGTA	972
QY	461	GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspLysPro---	479
Db	973	CAGCGCTGCGCAACGCTGAAAACCAAAACCGTGATTTCTCTCGACATCGAACGGGCG	1033
QY	480	-----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu	492
Db	1033	CGGAAATCTACGTCACCAAGAAATACACATCCGCGCAACACAAACCCGCGACGA---	1083
QY	493	AsnLeuValAlaAlaLysAlaArgHisIleuTyrAspMetProAspAsp-----	508
Db	1090	---GTCGTCCCGCGTGAATTCACGCAAAATGAATCCGACCTTACGACACACCTCCAAAGCTG	1144
QY	509	-----ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle	523

Db	1147	CAACGTTCCAAAGACCGCGTCGACCTTTTGCGGCTACTTCACAAATGTCACG-----	1197
Qy	524	LeuGIyargIIeSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspIuSer	543
Db	1198	-----TTTGATGCTGTCTCCCGCTTCCCGGACGCCCGACAAA---	1233
Qy	544	GIuSnGIuValIIeAspLeuProGIuArgThrAlaLeuAlaSnArgIuThrProAla	563
Db	1234	-----GTCGATTTG-----AACATGAGTCTGACCGAACGTTTCCACC-----	1265
Qy	564	AspValTYrGIuSerLysLysValProLeuTYrValPheValAlaSerAspIuProArg	583
Db	1269	-----	1265
Qy	584	AspGIyGIuIIeGIyLeu-----GIYrPGIySerAspThrGIyThrArgLeuValThr	601
Db	1270	---GGTTCCCTCGAATTGGAGCGCGGGTTGGGCTTCAAAGATACCGGG-----TTGGTCATG	1320
Qy	602	LysPheGIuIIeAsnLeuIIeAsnArgAspGIYrGIuAlaGIuValaGIuLeuArgLeu	621
Db	1321	TCCCGACGCGTTTCCCAAGAACACCTGTTCGGTACAGCGGCAAGTCGCGCCACGTGGCGCC	1380
Qy	622	SerGIuSerLysLysGIyValLysLeuTYrAlaThrLysProLeuSerIIeProLeuAsn	641
Db	1381	TCCAGAGCAAA-----ACCACGCTTACAGCGTCCGCTGTCTTACTACCCGCTCATTC	1434
Qy	642	AspGIuLeuArgAlaThrLeuGIYrGIuGIuGIuValaPheGIyIIeSerThrAsnGIy	661
Db	1435	ACGCGACAGCGGGGTCACCTGGGCTAC-----GATGTTACGGAAACGCTTCCACCCG	1488
Qy	662	PheAspLeuSerThrArgThrLeuGIuIIeAsnArgSerIIeIleGIuAsnGIy	681
Db	1489	CGCAAGACGATCGACACATCAAAACAATTAATAAACACACACGCA-----G--GAC	1536
Qy	682	GIYrPAsnArg-----ThTYrSerLeuArgTYrArgLeuAspLysLeu	696
Db	1537	CGAGCATCGCATAGACCTGGCTGTATCCGAATACGACCGCGGTGAATTCGGTTGGTG	1596
Qy	697	LysThrGIuAlaProProGIuThrTYrGIuAspLeuPro-----ValAspPheArg	713
Db	1597	GCAGAACACCTGACCGCTCAACACCTACACAAAGCGCCCAACACATACGCACTTATC	1655
Qy	714	Asn-----GIYrLys---ProSerGIuAlaIleuLeuAlaGIyValaIaValaIIeLys	730
Db	1657	AAGAAATACGGCAAAACCGACGCGACAGCGACGCTTCAAAAGCGCTGTACAAAGGT	1716
Qy	731	ThrVal-----AlaAspAsnLeuValaAsnProMetArgGIYrArg	744
Db	1717	ACCGTCGGCTGGGGCGCAACAAACCGACAGCGGCTTATGCGCGCGCGCTAC---	1773
Qy	745	GIuArgTYrSerLeuGIuValaGIySerSerGIyLeuValaSerAspAlaSnMetAlaIle	764
Db	1774	-----CTAGCGCGCGCTGAACCCCGCAATTC	1797
Qy	765	AlaArgAlaGIyIleSerGIyValTYrSerPheGIYAspAlaIaTYrGIySerAsnArg	784
Db	1798	GCCGCGCTGGCACCAACATCGCAATAC-----TACTCGCC	1833
Qy	785	AlaIIeGIuMetThrGIY-----	790
Db	1834	ACCCACAACCAACCACTGGTTCTTCCSCCTAGCAAAACCTTACGCTGATGCTCGCGGC	1893
Qy	791	-----GIYIleGIuAlaGIyTYrIleTYrSerAspAsnPheAsnIIeValaProTYrArg	808
Db	1894	GAACTCGGCATTTGGCGGCGGTAC-----GCGAGAACCAAAAGAAATCCCTCTT	1944
Qy	809	LeuArgPhePheAlaGIyGIuAspGIuSerIIeArgGIYTYrAlaIIeAspSerLeuSer	828
Db	1945	GAAACCTTCTACAGCGCGCGGCTGGGTGTGGGTGTGGGTGTGGGTGTGGGTGTGGGT	2004
Qy	829	ProIIeSerAspLysGIYTYrLeuThrGIyGIuGIuValaIleuAlaValaGIY-----	845


```

Db 2005 CCG-----AAAGTCATAGCAATACGGCAAAAAATACAGTACGGCGCAACAA 2055
Oy 846 ThrIaGluTyrAsnTyrGluPheMet-----Lysasp 856
Db 2056 AAACCCAAACGCTCCGGCCGACGCTCTTCCGATGCCCGCGCAAAAGACGGCGCAC 2115
Oy 857 LeuArGluAlaValPheGlyAspIleGlyAsnAlaTyrAsp-----870
Db 2116 GTCCCGCTGACCTGTTGGCCGACGACGAGCTGTGGACGCAAAACCTACGACGAC 2175
Oy 870 -----870
Db 2176 AACGACAGTTCGGGACCGCGGCAAGGTTCAAAACATTTAGCGCGCGCAATACCAT 2235
Oy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer 889
Db 2236 AAATTCACACCTTTACCAAGAAATTCGCGATTCGCCGCGCGCGGTTACCTGGCTCG 2295
Oy 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
Db 2296 CCTTTAGCCCGATGAATTCAGCTACGCGCTTACCCCGTGAAGAAAAACCGAAGACGAA 2355
Oy 908 ProLysLeuHisPhePheIleGlyThrProPhe 919
Db 2356 ATCCACAGCTTCCAATTCACATCGGACGACGATTC 2391

RESULT 32
AX202497 2394 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 10 from Patent WO0152885.
VERSION AX202497
KEYWORDS AX202497.1 GI:15392217
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis.
REFERENCE 1 (bases 1 to 2394)
AUTHORS Piazza,M.C., Rappuoli,R.C. and Giuliani,M.C.
TITLE Outer membrane vesicle (omv) vaccine comprising N. meningitidis
JOURNAL serogroup b outer membrane proteins
Chiron Spa (IT) Patent: WO 0152885-A 10 26-JUL-2001.
FEATURES
Source Location/Qualifiers
1..2394
/organism="Neisseria meningitidis"
BASE COUNT 655 a 758 c 586 g 395 t
ORIGIN
Alignment Scores:
Pred. No.: 2,29e-07 Length: 2394
Score: 259.00 Matches: 191
Percent Similarity: 33.48% Conservative: 121
Best local Similarity: 20.49% Mismatches: 342
Query Match: 5.48% Indels: 278
DB: Gaps: 40
US-09-914-168-2 (1-919) x AX202497 (1-2394)
Oy 116 ThrProLysLeuSerLeuGluGlnLeuPheAlaGlnLeuSerThrGluMetGlyIleAsnPro 135
Db 46 TCGCGCTTGGCACTTGGCCGACTTCACATCCAGACATCCGCTGACGCTTCAGCGT 105
Oy 136 AsnAsp-----TyrIleProGluTyrGlnGlyGlnGlnProAsnSer 149
Db 106 ACCGACCGGACTACCGATTCACCTACTGCGCTCAAGTCGGCGACACCTACAC---162
Oy 150 GluValAlaValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
Db 163 -----GACACACACGCGAGTGCATCATCAAAAGCCTGTAC 198
Oy 170 AlaArgLeuPheAsnAspGlyValAlaValAlaPro-----ArgLeuLysAlaLys 186

```

```

Db 199 GCCACCGGTTCTTTGACGAGCATCGCGTGCAGAACTCGGACGGCGAGCTCTGAC 258
Oy 187 PheTyrGlnSerSerGlnSerGlyGlnThrSerAlaIleGlySerSerHisGlnLysThr 206
Db 259 GTTATCGAAGCCCGCCACCATCGCTCGCTCAACATCACCGCGCAAAATGCTCAAAAC 318
Oy 207 GluProTyrAlaAsnIleLysAlaAlaLeuGlnLysPhe-----IleThrGlnLysSerAla 224
Db 319 GACGCC-----ATTAGAAAACCTCGAATTCGTTGGGCGTGGGCGAGTCCCAATAC 369
Oy 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln-----236
Db 370 TTTATACGGGACACATCAATACAGCAGTCCCGCTGAAAGAAATACCTCGGCGC 429
Oy 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTrpAsp 249
Db 430 GCGAACTCAATATTCAAATTCACGCCCAAACTAACCAAACTCGCCCAACCGCGTGCAC 489
Oy 250 IleAspLeuSerIleIleArgAsnSerIleGlyLysValAspValIleIleHisAspLeu 269
Db 490 ATCGACATCAGATTTGACGAGGGCAATCCGCCAAATACACGACATCGAATTGAA---546
Oy 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289
Db 547 GCGAACCAAGCTCTATTCGACCGCAACTGATCGCG-----582
Oy 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal-----306
Db 583 -----CAGATGCTGCTGACCGCAAGCGCGCATTTGGACATCG 618
Oy 307 -----PheHisGlyLysTyrGlnThrLysAsnLeuIleGlu 320
Db 619 CTGACAGCAACCAACCAATTCACAGAGAAATTTGCCCAAGACATGAAAAAGTAAAC 678
Oy 321 AsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgTrpLeuAspArgSerValAspVal 340
Db 679 GACTTCTACAGAACCAAGGCTACTGATTCGCGATCTCCGATACCGACATCAAAAC 738
Oy 341 IleLeuProAspAsnThrAlaAspValSerLeuLeuTyrAspTrpGlyThrGlnTyrArg 360
Db 739 AACGACAGCAAAACCAAGCAGACATCAAAATCACCGTCCAGAAAGCGGACGTTCCGT 798
Oy 361 PheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
Db 799 TGGGGCAAAATC-----TCCATGAAAGCGCACCAACGAA-----834
Oy 381 AspLysLeuProValLysArgGluLeuGlnGlnLeuLeuThrValAsnMetGlyGlu 400
Db 835 -----GTCCCAAAAGCCGAACCTGAAAAACCTGACCATGAAAGCCGCAAA 882
Oy 401 AlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
Db 883 TCGTACGACACCGCAGATGACCGCGTTTGGCT-----918
Oy 421 AsnMetValAsnThrGluIleValPheProGluArgGlnIleGlnAsnAspGluVal 440
Db 919 -----GAGATTGAGAACCGCATGGC 939
Oy 441 SerPheGlnIleSerSerSerArgTrpGluProAlaGlnValAspGluSerThrLeu 460
Db 940 TCGGACAGGCTACGCAATACG-----GAAATGAGCGTA 972
Oy 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro---479
Db 973 CAGCGGCTGCCAAAGCGCAAAACCAAAACGTCGATTTGCTCGTCGACATCGAACCGGCG 1032
Oy 480 -----IleGlnPheSerAlaSerAsnLeuIleGlnAspLysLeu 492
Db 1033 CGGAAATCTTACGTCACAGAAATCCACATCACCGCAACAAACCCGAGACGAA---1089
Oy 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrTrpAspMetProAspAsp-----508

```

Db 1090 ---GTGCTGCGCCGGCAATTCGCCAAATGCCGCTTACGACACCTCCAAAGCTG 1146
 Qy 509 -----ArgValIleuAlaIleAsnHisAspGlyValAsnArgSerIle 523
 Db 1147 CAAGCTCCAAAGACGGCTGACGCTTTGGGCTACTTCGACACGTCACG----- 1197
 Qy 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspIleuSer 543
 Db 1198 -----TTTGATGCGCGTCCGCGGCGGACACCGGACAA----- 1233
 Qy 544 GluAsnGluValIleAspLeuProGlyArgThrAlaLeuAlaAsnArgLysThrProAla 563
 Db 1234 -----GTGCAATTG-----AACATGACCTGACCGAAGCTTCACAC----- 1269
 Qy 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583
 Db 1269 ----- 1269
 Qy 584 AspGlyGlnIleGlyLeu-----GlyTrpGlySerAspThrGlyThrArgLeuValThr 601
 Db 1270 ---GGCTCGCTCGACTGAGCGGGCGGTGGAGACGATACCGGC-----CTGGTCATG 1320
 Qy 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlyLeuArgLeu 621
 Db 1321 TCCGCGAGCGTTTCCCAAGACAACTGTCGTACGCGCAAGTCGCGCGCTCGCGCCG 1380
 Qy 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
 Db 1381 TCAGAGAACAA-----ACCAAGCTCAACGGCTCGCTGCTTACGACCCGACTTC 1434
 Qy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661
 Db 1435 ACGGCAGACGGGGTACGCTGGGCTAC-----GATGTTACGGAAAGACCTTCACACCG 1488
 Qy 662 PheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGly 681
 Db 1489 CGCAAGACATGACCATCAGCATCAACATATTAACCCACGCGCA-----GGC 1536
 Qy 682 GlyTrpAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696
 Db 1537 GCAGGCAATCGCATGAGCGTGGCTGTACCGATACGACCGGTCGATTTGGGTG 1596
 Qy 697 LysThrGlnAlaProTrpGlyThrTrpGlnAspLeuPro-----ValAspPheVal 713
 Db 1597 GCAGAACACTGACCTGCACACCAACAAAGCGCCCAACACTATGCGCACTTATAC 1656
 Qy 714 Asn-----GlyLys-----ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730
 Db 1657 AAGAAATACGGCAAAACCGACGGCACAGACGGCACTTCAAGAGCTGCTGACAAAGGT 1716
 Qy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
 Db 1717 ACCGTGGCGCTGGGGCGGACAAACGACAGCGGCTTATGGCGGCGCGCTAC--- 1773
 Qy 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
 Db 1774 -----CTGACGGGCGGTGAACGCCCAATC 1797
 Qy 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
 Db 1798 GCCCTGCCCGGACGAACTGCATAC-----TACTCGCGC 1833
 Qy 785 AlaHisGlnMetThrGly----- 790
 Db 1834 ACCCAGACCAAACTGCTTCTCCCTTAAGCAAACTTCACGCTGATCGCGCGC 1893
 Qy 791 -----GlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArg 808
 Db 1894 GAAGTCGCGCATTTGCGCGGCTAC-----GGCAGAACCAAGAAATCCCTTCTTT 1944
 Qy 809 LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
 Db 1945 GAAAACTTCTACGGCGGCGCTGGTTCGTCGCGGATACGAAAGCGGCGCTCGGT 2004

Qy 829 ProIleSerAspLysGlyTyrIleThrGlnGlyGlnValIleuAlaValGly----- 845
 Db 2005 CCG-----AAAGTATGACGAATACGGCGGAAATAATAGCTACGCGGCAACAA 2055
 Qy 846 ThrAlaGlyTyrAsnTyrGlnPheMet-----LysAsp 856
 Db 2056 AAAGCCACAGTCCTCCGCGAGCTGCTTCCGATGCCCGGCGGAAAGACGCGCAC 2115
 Qy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
 Db 2116 GTCCGCTGAGCTGTGTCGCGACGACGACGCGTGGAGCGCAAAACCTACGAGAC 2175
 Qy 870 ----- 870
 Db 2176 AACAGCACTCCGCGACCGCGGCGGCTTCAAAACATTTACGCGCGCGCAATCCAT 2235
 Qy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer 889
 Db 2236 AAATCCACTTTTACCAACGAATTCGCTATTCGCCGGCGGCGGCTTACCTGCGCTCG 2295
 Qy 890 ProValGlyGlnValArgValAlaValAlaThrGlyValLysGlu-----GluLysAsn 907
 Db 2296 CCTTAAAGCCCGATGAATTCAGTACGCTTACCGCGTGAAGAAACCGGAAGACGAA 2355
 Qy 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
 Db 2356 ATCCAACGCTTCATTTCCAACTCGGACGACGCTTC 2391
 RESULT 33
 AE002375/c 13538 bp DNA linear BCT 25-MAY-2000
 LOCUS Neisseria meningitidis serogroup B strain MC58 section 17 of 206 of
 DEFINITION the complete genome.
 ACCESSION AE002375 AE002098
 VERSION AE002375.1 GI:7225394
 KEYWORDS
 SOURCE
 ORGANISM
 Neisseria meningitidis MC58.
 Neisseria meningitidis MC58
 Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
 Neisseria.
 REFERENCE
 1 (bases 1 to 13538)
 Achtel, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
 Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
 Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,
 Hickey, E.K., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,
 Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
 Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
 Qiu, H., Yamacheyan, J., Gill, J., Scarlato, V., Maignani, V.,
 Pizzia, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
 Rappuoli, R. and Venter, J.C.
 Complete genome sequence of *Neisseria meningitidis* serogroup B
 strain MC58
 JOURNAL Science 287 (5459), 1809-1815 (2000)
 MEDLINE 20175755
 PUBMED 10710307
 REFERENCE
 2 (bases 1 to 13538)
 Achtel, H., Saunders, N.J., Heidelberg, J., Jeffries, A.C.,
 Nelson, K.E., Eisen, J.A., Ketchum, K.A., Hood, D.W., Peden, J.F.,
 Dodson, R.J., Nelson, W.C., Gwinn, M.L., DeBoy, R., Peterson, J.D.,
 Hickey, E.K., Haft, D.H., Salzberg, S.L., White, O., Fleischmann, R.D.,
 Dougherty, B.A., Mason, T., Ciecko, A., Parksey, D.S., Blair, E.,
 Qiu, H., Yamacheyan, J., Gill, J., Scarlato, V., Maignani, V.,
 Pizzia, M., Grandi, G., Sun, L., Smith, H.O., Fraser, C.M., Moxon, E.R.,
 Rappuoli, R. and Venter, J.C.
 Direct Submission
 Submitted (17-MAR-2000) The Institute for Genomic Research, 9712
 Medical Center Dr, Rockville, MD 20850, USA
 TITLE
 JOURNAL
 FEATURES
 .source
 1. 13538
 /organism="Neisseria meningitidis MC58"
 /strain="MC58"

```
gene
CDS
/db_xref="taxon:122586"
/notes="serogroup: B"
complement(121..1377)
/gene="NMB0176"
complement(121..1377)
/gene="NMB0176"
/notes="similar to GB:102948 SP:p29011 PID:145703 GB:U00096
PID:1651591 percent identity: 76.98; identified by
sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="D-amino acid dehydrogenase, small subunit"
/protein_id="AAF40633.1"
/db_xref="GI:7225395"
/transl_table="MKVLYAGVAGVSSAMYLAEGEVTVIDRAGVAMETSPANA
GOLSYGTTTMAAGCIPITKALKMLFKSHPLPFPDGLSLOIEMQMIONCTAAHYO
TNEKRVRISEYSEKEMFRREAOFGNMEGKRGKGTLOIFROTREVEAKRDIALEHT
GVYPRRLKPECAEFEPALARVAKIAGCGLHLPADATGDCRLTENLYKLCQKGVAF
HFNQNTSRIDHNCIRKIVTETETGRFADAVCALGCFSTVLAQDLNLPITYPVGY
SLLPVYNSDGAIPSTVLDESIVVAITRFNRIKVGMAELSGYEIKLPEKRETTAL
VNDLPEEGDLSQASFMSGLRPMTPDPTPLIGTRFRDNLFLNTGHTLQMTSLGSA
KLVDIVSGKDTETSRDLSLRYOA"
complement(1456..2847)
/gene="NMB0177"
complement(1456..2847)
/gene="NMB0177"
/notes="similar to GB:AE000511 PID:2314080 percent
identity: 69.11; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="sodium/alanine symporter, putative"
/protein_id="AAF40634.1"
/db_xref="GI:7225396"
/transl_table="MOVFLDNPKAFPEFVSGWVSPIMLMLVGTILTVLLKLOF
TMVGYLKOAFMPKHKHSGEGHEDISHFALMTALSATIGTNAGVATLVVNGCP
GAVFMWMTAIFGMATKYGEGVALYKRVNNSKMGSGPMYIEKGLKNNKMAVA
FALFETFAFGSGSVSNVAOVOTSGFEPAYGILTVLTAAYVYGGIKGAKA
ASPIVPAWVYVYGLSLIATNSDMLPAMVLTIPSDASAQAVAGATGTVIRGVA
RCVSENEAGMSAPITAAAKDHPRLQALVYMTGFTDITVCSITGTLVWGLGA
GCEFYKPEVSGALITVTOKMLPGIGMIIVITGILFAYSTILGWCYGGCAVYVF
GEKFACTLRVGVSSVMTGLVSLDLWLASDTPFNLMLPMLIALMLAKIVNETR
DRQKITTNGELPH"
complement(3204..3980)
/gene="NMB0178"
complement(3204..3980)
/gene="NMB0178"
/notes="similar to PID:1718489 percent identity: 100.00;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="acyl - (acyl-carrier-protein) --UDP-N-
acetylglucosamine O-acyltransferase"
/protein_id="AAF40635.1"
/db_xref="GI:7225397"
/transl_table="MTLAIPTAVIDPKAELDSGVKGVATVIGPNOIGANTELGPHA
VINGTSIGENNRIFOFASLGEIPQDKRYRDEPTKLITGNGNTIREFTFPNGTGTGI
GETRIGDNTNIMAYCHLAHDCVIGNHTIFEANNASLAGHTIGDVTGLVYQFGR
IGDYMTNFAACVHKVDVDPYFMAFGSRAEPGLSEGRRNGFTAEQISAVKDYKTL
YHRGIPFEAKDILIRAEQALAVDFDFAQSARGLR"
complement(4067..4516)
/gene="NMB0179"
complement(4067..4516)
/gene="NMB0179"
/notes="similar to PID:1718488 percent identity: 100.00;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product=" (3R)-hydroxymyristoyl - (acyl carrier protein)
dehydratase"
/protein_id="AAF40636.1"
/db_xref="GI:7225398"
/transl_table="MDVQLPTEAKDIOKLIPHRPTQLDRITAFEPMKTLTAIKNVT
```

```
gene
CDS
/INPQEGHPDPLVMPGVLIIEAMAQCTLLASEGGRKEFFFGAIDEARFKR
QVLPQDLYEVEVELTSRIGIKGFENAVAKVDGQVAVAEATIMAKRRV"
complement(4450..5596)
/gene="NMB0180"
complement(4450..5596)
/gene="NMB0180"
/notes="similar to PID:1718487 percent identity: 100.00;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="UDP-3-O-(3-hydroxymyristoyl)-glucosamine
N-acyltransferase"
/protein_id="AAF40637.1"
/db_xref="GI:7225399"
/transl_table="MLPATYTLQITARLGMRGEDISVTAVRPLADAQAEHISFLA
NPKYKAEVHDSAGAVYSAKADGEGENLLIVADPPYLYFPAKVARLSPYKARGGI
HPYAVVEPATVPTSCIEIGANYIIGANTVIGSGCRILAAVVOHCKIGDEVLVAPNA
VYVGCTLGRVYEIHSAGAVIGADGELAFADDSKRIPTQAGVTTIGDDVEIGSNTNID
RGMSPPTTGNGTKIDNOYDIGHNCKIGSHYIAAKTGISGVTIGSTCITIGGVGTV
GHEIADKTTIGGCTSVTHSITESGKHLAGIFPMSTHEWARNNAVYIHRLEMMNRK
TIEQQLSDAGQDSK"
complement(5629..6129)
/gene="NMB0181"
complement(5629..6129)
/gene="NMB0181"
/notes="similar to GB:M34854 SP:p31519 PID:155479 percent
identity: 46.88; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="outer membrane protein OmpB, putative"
/protein_id="AAF40638.1"
/db_xref="GI:7225400"
/transl_table="VTRITRAPAALLIGCTACGAAHDTFOKIGINTERIYLESKQA
KRIQTTDSEFSARDELOQLQRELDLERLAEKLNMAKAAEEMKRLVAFRA
KQAEDEEYNNLRNEEFASLQONANRVITAKQEGXDVIQNNIYVTOYDVDSVI
KEMNR"
complement(6195..8588)
/gene="NMB0182"
complement(6195..8588)
/gene="NMB0182"
/notes="similar to GP:2460281 percent identity: 99.87;
identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="outer membrane protein Omp85"
/protein_id="AAF40639.1"
/db_xref="GI:7225401"
/transl_table="MKLQINSALMMLGISPLADFTQDIRVEGLQRTPESTVENY
LPAVGDITNDTHESALIKSLYATGFFDDVAVETADQQLLTVEPTIGSLNITGAK
MLQNDATIKRNLSEGLAOSQYFNQATLNOAVAGLEKYEYLGKLNIOITPKVRLAN
RVDIDITIDEKSAKITDIEFEGNOQVSDRKLKMSMLTGGITMTLIRSNONEKPF
AODMEKYVDFQNNNGYEDERLIDTIDTQDNEKTRQTKITVYHEGGRMKRVSIEGT
NEVPAEELKELTKMKPKKWEERQOMTVVLGEIQRMGSAGATASEISQPLPAKERT
VDVLIHTEPKKITYNEIHIITGNKKTIDEVVRRLRLROMESAPDTSLQSKRDELL
CYEPNVQDAVPLGLTDPDKVDLNNKLTDEVRSTGLSDSAGWODTGLWMSAGVQDMLF
TGSASALRSRSYTYLLNGSLSPFDPTFETADGVSLGADVKGAFDPKRAKTSJKOYT
TTAGAGLRMSVPVYEDYRVNVEGLVAEHLTYNTYAKARHVADELTKKTKGWDGDSFR
GWLKGTGVGWRNKTDLSALMPTREGYLGVNAELTAPSKLOXYSAHNNQMPPLPKST
FTLMLGGEVGIAGGYGRTKELPEPENVYGGCLASVKGESGCTGPKKYIDYIGEKISG
GNKKANVASLELPPMPCAKDARTYRUSLRFADAGSVWDGKTYDNNSSATGGRVQNTYG
AGNTHKSTFTNELHYSAGAVTWLSPILGPKFSAVPLPKKRPDEIQRFOFQIGTF"
complement(8645..9985)
/gene="NMB0183"
complement(8645..9985)
/gene="NMB0183"
/notes="conserved hypothetical protein; identified by
Glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF40640.1"
/db_xref="GI:7225402"
```


Oy	682	GLYTPRANAT	-----ThrtysSerLeuargtyrargleuaspIysLeu	696
Db	7052	GCAGGCATCCGAGTACGGCTCCGCTTACGGAATACGACGCGGTGAATTCGGTTGGTG		6992
Oy	697	LysThrGlnAlaProProGluThrTrpGlnAspLeuPro-----ValAspPheVal	713	
Db	6992	GCAGAACACGTCAGCCGTACACACTACACAAAGCGCCCAACACATATGCCACTTTATC		6933
Oy	714	Asn-----GlyLys---ProSerGlnAlaLeuLeuAlaGlyValAlaValHisLys	730	
Db	6932	AAGAAATACCGCCAAACCGACGCGACAGCGGAGCTTCAAGGCGCTGTACAAAGGT	6872	
Oy	731	ThrVal-----AlaAspAsnLeuValAsnProMetargGlytyrarg	744	
Db	6872	ACCGTCGGCTGGGGGGCCACMAAACCGACAGCGCTTATGGCGCAGCGCGCGCTAC	6818	
Oy	745	GlnArgTyrSerLeuGlnValGlySerSerGlyLeuValSerAspAlaMetAlaIle	764	
Db	6815	-----CTACGGGCGGTGAACCCCGAATC	6792	
Oy	765	AlaArgGlnGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnarg	784	
Db	6791	GCCCTGCGCTGGCAGCAAACTGCAATAC-----TACTCCGCG	6756	
Oy	785	AlaHisGlnMetProGly-----	790	
Db	6735	ACCCACACAAACAACTGCTTCTTCCCGCTGAGCAAAACCTTACAGCGTATGCTCCGGCGC	6696	
Oy	791	-----GlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrarg	808	
Db	6695	GAACTGCGCATTTGGCGGGCGCTAC-----GCGAGAACCAAGAAATCCCTTTCTTT	6645	
Oy	809	LeuArgPhePheAlaGlyAlaYasPrgInsPheIleArgGlyTyrAlaHisAspSerLeuSer	828	
Db	6644	GAAACTCTTACGCGCGCGCGCTCGCTGGCTGCGCGGTATCGAAAGCGCGACGCTCGGT	6585	
Oy	829	ProIleSerAspIysGlyTyrLeuPheThrGlnGlyGlnValLeuAlaValGly-----	845	
Db	6584	CCG-----AAAGCTTATGACGAATACGGCGGAAAAAATACGCTACGCGCGCAACAA	6534	
Oy	846	ThrAlaGlnTyrAsnTyrGluPheMet-----LysAsp	856	
Db	6533	AAACCAACGCTCCCGCCGAGCGTCTTCCGATGCCCGCGCGCAACGCGCGCAC	6474	
Oy	857	LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp-----	870	
Db	6473	GTCGCGCTGAGCGCTGTTGGCGACGCGAGCGAGTGTGGGACGCAAACTACGACGAC	6414	
Oy	870	-----	870	
Db	6413	AACAGCAGTTCGCGACGCGCGGCGAGGTTCAAAACATTTACGCGCGCGGCAATACCAT	6354	
Oy	871	---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer	889	
Db	6353	AAATGCAACGCTTACCAAAACGAATTCGCGCTATTCGCGCGCGCGCGGTACTCGTCTCG	6294	
Oy	890	ProValGlnGlnValAlaArgValAspValAlaThrGlnValLysGln-----GluGlyAsn	907	
Db	6293	CCTTTACGCGCGCGTGAATTCACGCTACGCGCTACCGCGGTGAAGAAACCGGAAAGACGAA	6244	
Oy	908	ProIleLysLeuHisPhePheIleGlyThrProPhe	919	
Db	6233	ATCCAAAGCGCTTCAATTCGAACCTGGCAGCGAGCTTC	6198	

RESULT 34
AC020884_1/c
WPCOMMENT

Fragment Name	Begin	End	Locus AC020884	Accession AC020884
AC020884_0	1	110000		
AC020884_1	100001	210000		
AC020884_2	200001	310000		

```

AC020884.3      300001      373038
Continuation (2 of 4) of AC020884 from base 100001 (AC020884 Mus musculus clone RP23-
Alignment Scores:
Pred. No.:      3.57e-05      Length:      110000
Score:          259.00      Matches:      120
Percent Similarity: 35.09%      Conservative:  86
Best local Similarity: 20.44%      Mismatches:   192
Query Match:    5.48%      Indels:      190
DB:             2      Gaps:        17
US-09-914-168-2 (1-1919) x AC020884_1 (1-110000)

```

QY	272	ProValTyrIleAspTyrArgAlaValAlaGluValArgGlyIleuGluAlaAspLysAla	291
Db	42289	CGGGCTTAATTGGGGGACCGAGTGTGTATTGGCGGGCGGGCGGACCGATTAAGAC	42230
QY	292	PheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisGlyLys	311
Db	42229	TATTGAAATTCGTCGATACTCGCCGCG---GCTATTGGACGGTACTGAACACAGCGCAT	42173
QY	312	TyrGluThrLysLysAsnLeuIleGluAsnIleSerAlaGluHisGlyTyrPheAspGly	331
Db	42172	TATGAAATTTCAAAAGACTCTTAACCAAGCATCGCTTGCGTAAAGCTTAATTGGATAGC	42113
QY	332	ArgThrLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeu	351
Db	42112	GAATTTACCAAGGACGACCTGGCATTTGCCGTGGCCCTGCCTTAAGCTTCTGGGATATT	42053
QY	352	IleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIleAspPro	371
Db	42052	GATTATTAACACTGGCGCAACGTTACCGCTTGGCGATGTGACCTTT	42008
QY	372	LysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGlu	391
Db	42007	-----GAAAGATCAACAAATCCCGCATTAACCTGTCAA	41975
QY	392	GlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSer	411
Db	41974	AATTTGGTGGCGTTTAAAGAGGCGGATGAGTACCAATCGAAAGATCTGGCGAAGTGAAC	41915
QY	412	AsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGlu	431
Db	41914	CGCGGACTTCTGCTACCGGTGGTGAATTAACCTGCGT-----GTGGTGGCTCCCAAA	41864
QY	432	ArgGluGlnIleGlnAsnAspGlnValSerPheGlnGlnSerSerSerArgThrGlu	451
Db	41863	-----TTTGATTAACGCGCGCAACGAAGAAGTATTA	41834
QY	452	ProIleGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAsp	471
Db	41833	CCA-----TTTGACG--	41825
QY	472	GlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLys	491
Db	41824	GGCGTGT-----GTTTGGCGCG--	41810
QY	492	LeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeu	511
Db	41810	-----	41810
QY	512	AlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGluArgIleSerAspAlaVal	531
Db	41810	-----	41810
QY	532	SerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuPro	551
Db	41809	-----CGAACGAAACACCATC--	41792
QY	552	GluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysVal	571
Db	41792	-----	41792

Oy 572 ProLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyr 591
 Db 41791 -----GAAACCGGGCGGTTAC 41774
 Oy 592 GlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAspLeuLysnArgAsp 611
 Db 41773 TCTACGGACGGTGGACCGCGGTGAACGACGCTGGAACAAACCGCTGGATGACCTTAT 41714
 Oy 612 GlyTyrGlnAlaGlyAlaGlyLeuArgLeuSerGluAspLysGlyValLysLeu--- 630
 Db 41713 GGTCAACACTGCACACACAGACTAGTATTTCCGGCCGGACACACACCTGCACCTTACAC 41654
 Oy 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
 Db 41653 TATAAATAGCGCTGCTGAATAATCCACTGCAACATATATTATTGTGACAGCGGCTTTT 41594
 Oy 651 GlnGlnGlnValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670
 Db 41593 AAGGCGACT-----GACCTGAACGATACCGAATCTGACTGCACCTACGCTG--- 41549
 Oy 671 HisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrPasnArgThrTyrSerLeuArg 690
 Db 41548 ---GTGGCTTCTCGTACTGGGATCTCTCCACGGCGTGGACAGCTGCCTTAACCTCGCC 41492
 Oy 691 TyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProVal 710
 Db 41491 TGGAGTCTCGAC----- 41480
 Oy 711 AspPheValAsnGlyLysProSerGlnGlnAlaLeu-----AlaGlyValAlaVal 728
 Db 41479 CACTTTACTGAGTGAATTAATCAATACACAGATGCTGTTTATCTGGGCTGATGATT 41420
 Oy 729 HisLysThrValAlaAspAsnLeuValAsnProMetArgLysTyrArgGlnArgTyrSer 748
 Db 41419 AGCGGACCGCTTCTCGTGGTGGCTGATGCCAACCTGGGCGACTCGCAACGCTACTCT 41360
 Oy 749 LeuGlnValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGly 768
 Db 41359 ATCACTACTCTCAACACCGCTGGGGTTCAGATGTCGATTTCTCCGTTTCCAGCGCAG 41300
 Oy 769 IleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMet 788
 Db 41299 AACCTCTGATCCGACACTGATGATGC-----CATGCTTTT 41261
 Oy 789 ThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAspAsnHisValProTyrArg 808
 Db 41260 GTTACACGGCGACCGCTGGCTGATGAACCGGCTGATTCGACAACTACCCCGGAT 41201
 Oy 809 LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
 Db 41200 CTGGCTTTCTTCTCGCGGGGCGACCGCAGTATTCGTGGTCAATAACAA-TCTATCGCT 41142
 Oy 829 ProIleSerAspLysGlyTyrLeuThrGly-----GlyGlnValLeuAlaVal 844
 Db 41141 CCGAAATACCGCCAGCGGTGACTGAAGGGCCTCGAAGTGATTAACCGGATC----- 41091
 Oy 845 GlyThrAlaGlyTyrAsnTyr 851
 Db 41090 -----GCTGAATACCACTAC 41076
 RESULT 35
 NMA12491
 LOCUS NMA12491 340806 bp DNA linear BCT 04-DEC-2000
 DEFINITION Neisseria meningitidis serogroup A strain Z2491 complete genome:
 segment 1/7.
 ACCESSION AL162752
 VERSION AL162752.2 GI:7378778
 KEYWORDS
 ORGANISM
 SOURCE Neisseria meningitidis Z2491.
 Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
 Neisseria.
 REFERENCE
 1 (bases 1 to 340806)

AUTHORS Parkhill,J., Achtman,M., James,K.D., Bentley,S.D., Churcher,C., Klee,S.R., Morelli,G., Basham,D., Brown,D., Chillingworth,T., Davies,R.M., Davis,P., Devlin,K., Felwell,T., Hamlin,N., Holroyd,S., Jagsels,K., Leather,S., Moule,S., Mungall,K., O'Neill,M.A., Rajandream,M.A., Rutherford,K.W., Skellerns,M., Skellerns,J., Whitehead,S., Spratt,B.G. and Barrall,B.G.
 TITLE Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
 JOURNAL Nature 404 (6777), 502-506 (2000)
 MEDLINE 20222556
 PUBMED 10761919
 REFERENCE 2 (bases 1 to 340806)
 AUTHORS Parkhill,J.
 TITLE Direct Submission
 JOURNAL Submitted (30-MAR-2000) Submitted on behalf of the Neisseria sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
 COMMENT Notes:
 Details of N. meningitidis sequencing at the Sanger Centre are available on the World Wide Web.
 (URL, http://www.sanger.ac.uk/Projects/N_meningitidis/).
 FEATURES
 source location/Qualifiers
 1..340806
 /organism="Neisseria meningitidis Z2491"
 /strain="Z2491"
 /db_xref="taxon:122587"
 /note="serogroup: A"
 complement(38..358)
 /gene="NMA0001"
 complement(38..358)
 /note="NMA0001, unknown, len: 106 aa"
 /gene="NMA0001"
 /codon_start=1
 /transl_table=11
 /product="hypothetical protein NMA0001"
 /protein_id="CAB83521.1"
 /db_xref="GI:7378779"
 /db_xref="SPTREMBL:O94X93"
 /translation="MSNTQITRSADPTSRAGALDIANNNGTTRVLRHTDQPYKWHYNDGEVFAVMCEVDVMHYRENEEHLVRLKSDIDFYAGTGEVHVRPRETRILVIEKBCSV"
 complement(414..2438)
 /gene="nuoL"
 complement(414..2438)
 /gene="nuoL"
 /EC_number="1.6.5.3"
 /note="NMA0002, nuoL, NADH dehydrogenase I chain L, len: 674 aa; similar to many e.g. NUOL_RHOCA NADH dehydrogenase I chain L (EC 1.6.5.3) (700 aa). Contains Pfam match to entry PF00361 oxidored_g1, NADH-ubiquinone/plastoquinone (complex I), various chains, and to entry PF00662 oxidored_g1_N, NADH-ubiquinone oxidoreductase (complex I), chain 5 N-terminus"
 /codon_start=1
 /transl_table=11
 /product="NADH dehydrogenase I chain L"
 /protein_id="CAB83322.1"
 /db_xref="GI:7378780"
 /db_xref="SWISS-PROT:O9JX92"
 /translation="NDMTLYLIALVPLAGSLAGLEGKIGRAGAHVTVTLGAVS AVLSAVYLMGLGSRKAFDENVTYMLMGGLDFSVGLVDNMTAMMYVAVGSLAV HIYITGYNHDEKVGVOREFSYISLFTFSMLIMSNTFOLFPGWEAVGLVLLIGF YFKRSPHNEFANLAKPLINVGDEGFLGLGLVAVFGSLRYODVAYLPVONNATIO LEPGVENSLITYTCILLPGAMKSNAPFLHWLPDSMGSPDISALHIAATMVATG FVMSRSPITEMSSTALSYIYIGALITLAPKFLVGIYONDIKRVAYSTLSLQWTV ALGASAVYAMHVHTAEFFKALPLEAAGSALITGHHDDMRKCNLKYMTWLTW TL LIGNLSLGTPEFSGFYSDSILTEAKYSTLPGSGFAVAVAVFAVFAFRQFM VFNHGEKRSLEPHNSDGHSEHGLGKNDNHSPLVYTLPLILAVSVYIIGYAI EPMLYGDFEKDVIYFNADAPHPYFENKYIYDALFYNVFAKGTALGTFPKVGDALID NGIVNGSRIVGAVAAYVRKVDGTFTYTAAMVGVLLDMTETGLR"
 1035..1044
 /note="Core DNA uptake sequence: gccgtctgaa"

misc_feature

```

misc_feature      /label=DUS
                  complement(1116..2030)
                  /gene="nuoL"
                  /note="Pfam match to entry PF00361 oxidored_g1,
                  NADH-ubiquinone/plastoquinone (complex I), various chains,
                  score 351.80, E-value 7.3e-102"
                  complement(2061..2243)
                  /gene="nuoL"
                  /note="Pfam match to entry PF00662 oxidored_g1_N,
                  NADH-ubiquinone oxidoreductase (complex I), chain 5
                  N-terminus, score 77.20, E-value 3.5e-19"
                  complement(2487..2828)
                  /gene="NMA0003"
                  complement(2487..2828)
                  /gene="NMA0003"
                  /note="NMA0003, unknown, len: 113 aa"
                  /codon_start=1
                  /transl_table=11
                  /product="hypothetical protein NMA0003"
                  /protein_id="CAB83323.1"
                  /db_xref="GI:7378781"
                  /db_xref="SPTREMBL:Q9JX91"
                  /translation="MMDKNOLEDFHKLINLYOEALNPQPYKATRFQIYVEFGK
                  EADKLSTGEKKTGTGTETLISGGVHALKYSMEYLVLQKPCDLFTBQLAVARK
                  RLERYGVSPK"
                  complement(2833..2837)
                  complement(2952..3527)
                  /gene="NMA0004"
                  complement(2952..3527)
                  /gene="NMA0004"
                  /note="NMA0004, unknown, len: 191 aa; similar to
                  hypothetical proteins e.g. Y977_HAEIN HI0977 (191 aa),
                  fasta scores: E(): 0: 80.5% identity in 190 aa overlap,
                  and to proteins involved in cell cycle e.g. TR:054679
                  (EMBL:AF036487) Laccococcus lactis plasmid pNZ4000
                  putative mobilization protein (200 aa), fasta scores: E():
                  0, 55.1% identity in 167 aa overlap, and FIC ECO11 cell
                  filamentation protein FIC (200 aa), fasta scores: E():
                  0.093, 29.4% identity in 126 aa overlap"
                  /codon_start=1
                  /transl_table=11
                  /product="hypothetical protein NMA0004"
                  /protein_id="CAB83324.1"
                  /db_xref="GI:7378782"
                  /db_xref="SPTREMBL:Q9J0R9"
                  /translation="MSENPRIKTKMSIDIOSLHNRARFESGDIDRIVGTAGIQO
                  IRYLFGGLYDAFOIREDNISGGRFANMAYLKAIKIDMPERTFEETIAYVE
                  MNIAHPLEGNRSTRIMDLVKKLVVMQNVSKTLVLAEMERSPVNDLEHLEFL
                  LNDNLTDVDNKEIIFKGIOSYIEGYEKG"
                  complement(3516..3525)
                  /gene="NMA0004"
                  /note="Core DNA uptake sequence: gccgctctgaa"
                  /label=DUS
                  complement(3561..3866)
                  /gene="nuoK"
                  complement(3561..3866)
                  /gene="nuoK"
                  /EC_number="1.6.5.3"
                  /note="NMA0005, nuoK, NADH dehydrogenase I chain K, len:
                  101 aa; similar to many e.g. NUOK_RHOCA NADH dehydrogenase
                  I chain K (EC 1.6.5.3) (102 aa), fasta scores: E():
                  1.1e+24, 65.0% identity in 100 aa overlap. Contains Pfam
                  match to entry PF00420 oxidored_g2,
                  NADH-ubiquinone/plastoquinone oxidoreductase chain 4L"
                  /codon_start=1
                  /transl_table=11
                  /product="NADH dehydrogenase I chain K"
                  /protein_id="CAB83325.1"
                  /db_xref="GI:7378783"
                  /db_xref="SPTREMBL:Q9J0Q9"
                  /translation="MTLTHYLVGLALFGISAMGIFMNRKNVLLMSTIEMLAVN
                  FNFIAFSQHLGDTGQIFVFPVLVVAASAIIGLAIVLVYRNQTTINVADDELKG"
                  complement(3564..3863)

```

```

gene              /gene="nuoK"
                  /note="Pfam match to entry PF00420 oxidored_g2,
                  NADH-ubiquinone/plastoquinone oxidoreductase chain 4L,
                  score 141.50, E-value 1.5e-38"
                  complement(3863..4534)
                  /gene="nuoJ"
                  complement(3863..4534)
                  /gene="nuoJ"
                  /EC_number="1.6.5.3"
                  /note="NMA0006, nuoJ, NADH dehydrogenase I chain J, len:
                  223 aa; similar to many e.g. NUOK_RHOCA NADH dehydrogenase
                  I chain J (EC 1.6.5.3) (202 aa), fasta scores: E():
                  9.7e+21, 41.3% identity in 201 aa overlap. Contains Pfam
                  match to entry PF00499 oxidored_g3,
                  NADH-ubiquinone/plastoquinone oxidoreductase chain 6"
                  /codon_start=1
                  /transl_table=11
                  /product="NADH dehydrogenase I chain J"
                  /protein_id="CAB83326.1"
                  /db_xref="GI:7378784"
                  /db_xref="SPTREMBL:Q9JX90"
                  /translation="MTFOLLIFXIFAVIILYGAIKTVAKNPVNAHLIVTPCSAM
                  LMLMQAEFGVTLVVYGVAVVLFVVMYMNIDIEBRAGFRNARVAGVGTLL
                  AVALLILVNPRTDLAAGLMKIDIPADYNNIRDLGRITDILPELAAVLLGLMV
                  AATVLRKTVNPKRMDPADQVVRADQGRMLVKMAVKKPQTESAESEVSDDLKPK
                  EESKA"
                  complement(3874..3879)
                  /gene="nuoJ"
                  complement(4034..4525)
                  /gene="nuoJ"
                  /note="Pfam match to entry PF00499 oxidored_g3,
                  NADH-ubiquinone/plastoquinone oxidoreductase chain 6,
                  score 77.90, E-value 2e-19"
                  complement(4570..5277)
                  /gene="NMA0007"
                  complement(4570..5277)
                  /gene="NMA0007"
                  /note="NMA0007, unknown, len: 235 aa"
                  /codon_start=1
                  /transl_table=11
                  /product="hypothetical protein NMA0007"
                  /protein_id="CAB83327.1"
                  /db_xref="GI:7378785"

RBS
complement(3874..3879)

misc_feature
complement(4034..4525)

gene
complement(4570..5277)

CDS
complement(4570..5277)

Alignment Scores:
Pred. No.: 0.000159
Score: 259.00
Percent Similarity: 33.488
Best Local Similarity: 20.498
Query Match: 5.488
DB: 1
Gaps: 40

US-09-914-168-2 (1-919) x NMA122491 (1-340806)
QY 116 ThrProLeuSerLeuGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
Db 84027 TCGCCTTGCGACTTGCCGACTTCACATCCAAAGCATCCCGGCGAAGCGTTGACGGT 84086
QY 136 AsnAsp-----TyrIleProGluTyrGlnGlyGlnGlnProAsnSer 149
Db 84087 ACCGAGCCGAGACGCTATTCACATGCTGCCCGCAAAAGTCGGGACACCTACAACT--- 84143
QY 150 GluValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
Db 84144 -----GACACACAGCGCAATGCCATCATCAAAAGCTGTGAC 84179
QY 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
Db 84180 GCCACCGGTTCTTTCAGAGCTAGCGGCTCGAAACTGCGGAGCGGAGCTCTGTCGAC 84239
QY 187 PheTyrGlnSerSerGlnSerGlyGlnThrSerAlaIleGlySerSerHisGlnLysThr 206
Db 84240 GTTATCGAAGCCGCCACATGGCTCGCTCAACATGACCGGCGCAAAAGTGTGCAAAAC 84299

```


QY 207 GluProTyrAlaAsnIleLysAlaLeuGluAsp-----IleThrGlnIleuSerAla 224
 Db 84300 GAGCCC-----ATTAAAGAAAACCTCGAATCGTTCGGGCTGGCCAGTCGCAATAC 84350
 QY 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgIn----- 236
 Db 84351 TTTAAATGAGCGCAGCATCAATCAGCATCGCCGGCTGAAGAAAGAAATACCTCGGGCGC 84410
 QY 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
 Db 84411 GGCAAACTCAATATCCAAATACGCGCCAAAGTACCAACTGCGCCGACCGCTGCAC 84470
 QY 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
 Db 84471 ATGCACATCAAGATTGAGAGAGGCAATCCGCCAAATACACCGACATCGAATTGAA---- 84527
 QY 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289
 Db 84528 GGCACCAAGCTCTATTCCGCGCAACTGATGCGG----- 84563
 QY 290 LysAlaIleThrIleThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
 Db 84564 -----CAGATGTCGCTGACCGAAGCGCGCATTTGGACATGG 84599
 QY 307 -----PheHisIleGlyTyrTyrGluThrLysLysAsnLeuIleGlu 320
 Db 84600 CTGACAGCAAGCAACCAATTCACGACAGCAAAATTTGCCCAAGACATGGAAAAAGTAAAC 84659
 QY 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPheAspArgSerValAspVal 340
 Db 84660 GACTTCTACCGAAGACAGCGCTACTTCCGATTCGCGATCTCGATACCGATCCCAACC 84719
 QY 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArg 360
 Db 84720 AACCAAGCAAAACCAAGCAGACATCAAAATACCGTCCAGGAAGCGGACGTTTCGGT 84779
 QY 361 PheAspGluValAlaPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
 Db 84780 TGGGCGAAAGTC-----TCCATCGAAGCGCACCAACGAA----- 84815
 QY 381 AspLysLeuProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMetGlyLeu 400
 Db 84816 -----GTCGCCAAAGCGCAACTGGAAAAAATCTGACCATCAAGCCCGCAAA 84863
 QY 401 AlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
 Db 84864 TGTACGAAAGCCAGACAGATACCGCGCTTTGGGT----- 84899
 QY 421 AsnMetValAsnThrGluIleValIlePheProGluArgGlnIleGlnAsnAspGluVal 440
 Db 84900 -----GAGATTCAGAACCCGCAATGGGC 84920
 QY 441 SerPheGlnGlnSerSerSerSerArgThrGluProAlaGlnValAspGlnSerThrLeu 460
 Db 84921 TCGCGAGGCTACGATACGC-----GAATTCAGCGTA 84953
 QY 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro--- 479
 Db 84954 CAGCGCGTGGCCAAAGCGCGAAACCAAAACCGTCGATTCGCTCTGCATCAACCGGCGC 85013
 QY 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492
 Db 85014 CGGAAATATCTACGTCAACGAATATCCACATCAACCGCACACAAACACCGCGACGAA--- 85070
 QY 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp----- 508
 Db 85071 ---GTCGTGCGCCCGCAATTCGCGCAATCCGCGCTTACCGATCAACCTCCAAAGCTG 85127
 QY 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAlaAsnArgSerIle 523
 Db 85128 CAACGCTCCAAAGAGCGGCTGAGCTTTTGGGCTACTTCGACAGACGTACAG----- 85178

QY 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543
 Db 85179 -----TTTGATGCCGTCGCCCTTGGCGGACACCGCCGCAAA---- 85214
 QY 544 GluAsnGluValIleAspLeuProGluArgThrAlaAlaLeuAlaAsnArgLysThrProAla 563
 Db 85215 -----GTCCGATTTC-----AACATGAGCTGCACGAAAGCTTCCACCC----- 85250
 QY 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583
 Db 85250 ----- 85250
 QY 584 AspGlyGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr 601
 Db 85251 ---GGCTGCTGACTTGAAGCGCGGCTGGGTACAGATATACCGC-----CTGGTCATG 85301
 QY 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaLeuLeuArgLeu 621
 Db 85302 TCCGCGAGCGGTTCCCAAGACAACTGTTCGGTACGCGCAAGTCGCGCCCTGCGCGCGC 85361
 QY 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
 Db 85362 TCACGAGGCAAA-----ACCAAGCTCAACAGCGCTGCTGTTTACGACCCGCTACTTC 85415
 QY 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661
 Db 85416 ACGGAGACGGGGTCAAGCTTGCGGTAC-----GAGCTTACGAAACCTTCGACCGG 85469
 QY 662 PheAspLeuSerThrArgThrLeuGlnHisGlyIleSerArgSerIleIleGlnAsnGly 681
 Db 85470 CCGAAAGCATCGACAGCATCAAAACATATAAACCCACACGCA-----CGC 85517
 QY 682 GlyTyrAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696
 Db 85518 GCAGGCATCCGATGACGTCGCTGTTACCGAATACGACCGCGTGATTTCCGTTTGGTG 85577
 QY 697 LysThrGlnAlaProProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713
 Db 85578 GCAGAAACCTCGACCGGTCAACACGTACAAACAGCCCAACACATACGCGCTTATTC 85637
 QY 714 Asn-----GlyLys-----ProSerGlnGluAlaLeuLeuAlaGlyAlaValHisLys 730
 Db 85638 AAGAAATACGGGCAAAACCGACGCGACAGCGCAGCTTCAAAAGCGCTGTACAAAGCT 85697
 QY 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
 Db 85698 ACCGTGCGGTGGGGCGGCACAAACCGACAGCGGCTTATGCGCGACGCGGCTAC--- 85754
 QY 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
 Db 85755 -----CTGACGGCGGTGAACCGCGCAATTC 85778
 QY 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
 Db 85779 GCCCTGCGCGCGCGCAACATGCAATAC-----TACTCCGCC 85814
 QY 785 AlaHisGlnMetThrGly----- 790
 Db 85815 ACCCAACACCAAACTGGTTTCCCTTAAGCAAAACCTTACCGCTGATGCTGGCGGC 85874
 QY 791 -----GlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArg 808
 Db 85875 GAAGTCGCGCATTCGCGGCGCGTAC-----GCGCAAGCAACCAAAATATCCCTTCTTT 85925
 QY 809 LeuArgPhePheAlaGlyLysAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
 Db 85926 GAAACCTTCTACGCGCGCGCGCTGGGTGCGCGGATACGAAAGCGGACAGCTCGCT 85985
 QY 829 ProIleSerAspLysGlyTyrLeuThrGlyGlnValLeuAlaValGly----- 845
 Db 85986 CCG-----AAAGTATATGACGAATACGCGGAAATAATCAGCTACGCGGCAACAA 86036
 QY 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856

Db 86037 AAGGCAACGTCTCCGCGAGCTGCTTCCGATGCCGGCGGCAAGACGGCGCAC 86096
Oy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
Db 86097 GTCCGCTGAGCGCTTTGGCCGACGACGACGCTGTGGAGCGCAAAACCTACGACGAC 86156
Oy 870 ----- 870
Db 86157 AACACGACTCCGCGACGCGCGGAGGTTCAAAATTTACGCGCGCGCAATACCAT 86216
Oy 871 ---LysGlyPheThrAsnAspRhrIleGlyAlaGlyValAlaArgTrrAlaSer 889
Db 86217 AAATCCACTTTACCAAGCAATTCGGCTATTCGCGCGCGCGCTTACCTGGCTTCG 86276
Oy 890 ProValIleGlyIleValAlaArgValAlaSerValAlaThrGlyValGly-----GluGlyAsn 907
Db 86277 CCTTACGCGCGGATGAATTCAGCTACGCTACCGCTGAGCAAAACCGGAGACGAA 86336
Oy 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
Db 86337 ATCCAAACCTTCCAAATTCGACCTCGGACGACGCTTC 86372

RESULT 36
AX043922/c AX043922 349980 bp DNA linear PAT 24-NOV-2000
LOCUS Sequence 1 from Patent WO0066791.
DEFINITION AX043922
VERSION AX043922.1 GI:11342850
KEYWORDS
SOURCE Neisseria meningitidis.
ORGANISM Neisseria meningitidis.
Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
Neisseria.

REFERENCE 1 (bases 1 to 349980)
AUTHORS Pizze, M., Hickey, E., Peterson, J., Rettlein, H., Venter, J. C.,
Mastigiani, V., Galeotti, C., Mora, M., Ratti, G., Scarselli, M.,
Scariello, V., Rappunli, R., Frazer, C. M., and Grandi, G.
TITLE Neisseria genomic sequences and methods of their use
JOURNAL Patent: WO 0066791-A 1 09-NOV-2000;
CHIRON CORPORATION (US); THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES
source location/Qualifiers
1..349980
/organism="Neisseria meningitidis"
/db_xref="taxon:487"
/note="sequence too long, cut in 8 pieces.-seq 1: 1 to
349980 349980 bases--seq 108: 300001 to 649980 349980
bases--seq 109: 600001 to 949980 349980 bases--seq 110:
900001 to 1249980 349980 bases--seq 111: 120001 to 1549980
349980 bases--seq 112: 1500001 to 1849980 349980 bases--seq
113: 1800001 to 2149980 349980 bases--seq 114: 2100001 to
2273325 172325 bases"

BASE COUNT 83241 a 85091 c 95206 g 86442 t
ORIGIN

Alignment Scores:
Pred. No.: 0.000164 Length: 349980
Score: 259.00 Matches: 191
Percent Similarity: 33.48% Conservative: 121
Best Local Similarity: 20.49% Mismatches: 342
Query Match: 5.48% Indels: 278
DB: 6 Gaps: 40

US-09-914-168-2 (1-919) x AX043922 (1-349980)
Oy 116 ThrProLeuSerLeuGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
Db 180761 TCGCCTTGGGACTTCCGACATTCACACATCCGCGCTGAGGCTTGAGGCT 180702
Oy 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149
Db 180701 ACCGAGCGGACGATTCACACTACCTGCGCCCTCAAAAGTCGGGACACCTACACAC--- 180645

Oy 150 GluValValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
Db 180644 -----GACACACACGCGAGCTGCATCATCAAAAGCGCTGAC 180609
Oy 170 AlaArgLeuPheAsnAspGlyValAlaAsnLysValPro-----ArgLeuLysAlaLys 186
Db 180608 GCCACCGGCTTCTTGGAGAGCATGCGCGTGCAGAACTCGCGGACGCGAGCTCTGTGACC 180549
Oy 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
Db 180548 GTTATCGAACGCCGCCACCATGGCTCGCTCAACATCGACGCGCGCAAAATCTCTCAAAAC 180489
Oy 207 GluProTyrAlaAsnIleLysAlaAlaLeuGluAsp-----IleThrGlnGluSerAla 224
Db 180488 GACGCG-----ATTAAAGAAACCTCGAATCGTTGCGGCTGCGCGCATCGCAATAC 180438
Oy 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
Db 180437 TTTAATCAGCGGACACATCAATGAGGACGTCGCCGCTGAAAGAAAGATACCTCGGCGC 180378
Oy 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
Db 180377 GCGAAACTCATATATCCAAATGACGCGCGCAAAATGTAACCAACCTCGCGCGACCGGCTGAC 180318
Oy 250 IleAspLeuSerIleIleArgAsnSerIleGlyLeuValAspValIleIleHisAspLeu 269
Db 180317 ATCCGACATCAGATTTGACAGGAGGCAAAATCCGCAAAATACCGCAATTCGAAATTTGAA--- 180261
Oy 270 GlyGluProValTyrIleAspTyrArgValAlaValGluValArgGlyGlyAlaAspAsp 289
Db 180260 GGCACCAAGCTCTATTCGCGCGCAAACTGATCGG----- 180225
Oy 290 LysAlaPheThrThrValAlaAlaAspGluValProLeuLeuIleGlyAspVal----- 306
Db 180224 -----CAAAATGCTCCGACCGCAAGCGCGGATTTGGACATG 180189
Oy 307 -----PheHisIleGlyLysTyrGlyGluThrLysLysAsnLeuIleGlu 320
Db 180188 CTGACACGACGAACCAATTAACGAGCAGCAAAATTTGCCAAGATATGAAAAAGTAAACC 180129
Oy 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrrPheAsnArgSerValAspVal 340
Db 180128 GACTTTCACCAAAATTAACGCTACTTCGATTTCCGTCATCCGATACCGACATCAAAACC 180069
Oy 341 IleLeuProAspAsnThrAlaAspValAspSerLeuIleTyrAspThrGlnTyrArg 360
Db 180068 AACGACGACAAACCAACGACGACATCAAAATTCACCGTCACGAGGCGGACGTTCCGT 180009
Oy 361 PheAspGlyValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
Db 180008 TGGGGCAAAAGTC-----TCCATCGAAGCGGACACCAACGAA----- 179973
Oy 381 AspLysLeuProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMetGlyGlu 400
Db 179972 -----GTCCCAAAAGCGGACCTGGAACAAACCTGACCATGAAACCGCGGCAAA 179925
Oy 401 AlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
Db 179924 TGCTACGAAACCCAGAGATGACCGCGCTTTGGCT----- 179889
Oy 421 AsnMetValAsnThrGlnIleValPheProGluArgGlnIleGlnAsnAspGlnVal 440
Db 179888 -----GAGATTGCAAAACCGATGCGC 179868
Oy 441 SerPheGlnGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu 460
Db 179867 TCGGACGAGCTACGATACAG-----GAAATACGGCTA 179835
Oy 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro--- 479
Db 179834 CAGCGCGTGGCGAAGCGTGAACCAAAACGTCGATTTGCTCGACATCGAAGCGGCGC 179775
Oy 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492

Db 179774 CGGAAATCTAGCTCAACGAATACACATCTACCGGCAACAAACCAACCGCGACGAA--- 179718
Oy 493 AsnLeuValAlaAlaLysAlaArgHisLeuTYrAspMetProAspArp----- 508
Db 179717 ---GTCGTCGCGCGCTGAATTCGGCCAAATGGAAATCCGACCTTTCGACACCTCCAACTG 179661
Oy 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle 523
Db 179660 CAACGTTCCAAAGACGGCGCTGACACTTTTGGCGTACTTCGCAATGTCCAG----- 179610
Oy 534 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGlySer 543
Db 179609 -----TTTGATGCTGTCTCCCGCTGGCGCGACCGCCGCAAAA--- 179574
Oy 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
Db 179573 ---GTCGATTTG---AACATGAGTCTGACCCGACGCTTCCACC----- 179538
Oy 564 AspValTYrGlnSerLysLysValProLeuTYrAlaIleValAlaSerAspLysProArg 583
Db 179538 ----- 179538
Oy 584 AspGlyGlnIleGlyLeu-----GlyTrpGlySerAspThrGlyThrArgLeuValThr 601
Db 179537 ---GGTTCCTCCGATTTTATGCGCGGGTGGGTTCGAATATCCGG-----TTGGTCAAG 179487
Oy 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTYrGlnAlaGlyValAlaGluLeuArgLeu 621
Db 179486 TCCCGAGCGCGTTTCCCAAGACAACTGTGCGTACGGGCAAGTCGGCGGCACTCCGGCC 179427
Oy 632 SerGluAspLysLysGlyValLysLeuTYrAlaThrLysProLeuSerHisProLeuAsn 641
Db 179426 TCCAGAGCAAAA-----ACCACCTTAAACGCGTCCGTGCTTACTGACCCGCTACTTC 179373
Oy 642 AspGlnLeuArgAlaThrLeuGlyTYrGlnGlnGluValPheGlyHisSerThrAsnGly 661
Db 179372 ACGCGACAGCGGGGTCAGCGCTGGCTAC-----GATGTTTACGAAAAAGCTTCGACCCG 179319
Oy 662 PheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGly 681
Db 179318 CGCAAGCATGACGACCGCATCAACAAATATGAAACACCAACGGCA-----GGC 179271
Oy 682 GlyTrpAsnArg-----ThyTrpSerLeuArgTYrArgLeuAspLysLeu 696
Db 179270 GCAGCATCCCGCATGAGCGCTGCTTACCGAATAACAGCCGCGAATTCGCTTGGTG 179211
Oy 697 LysThrGlnAlaProProGluThrThrProLysAspPro-----ValAspPheVal 713
Db 179210 GCACAAACACTGACCGCTCAACACCTCAACAAAGCGCCAAACACTATGCGCACTTATC 179151
Oy 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730
Db 179150 AAGAAATACGGCCAAACCGACGACACACCGGACGACTTCAAAAGGCTGCTACAAAGCT 179091
Oy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTYrArg 744
Db 179090 ACCGTCGGCTGGGGCGGCAACAAACGACGCGCGTTATGCGCCGACGCGCGCTAC--- 179033
Oy 745 GlnArgTYrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
Db 179033 ---CTGACGGCGCTGCAACGCGCGAATC 179010
Oy 765 AlaArgAlaGlyIleSerGlyValTYrSerPheGlyAspAsnAlaTYrGlySerAsnArg 784
Db 179009 GCCCTGCTGGACGAACACTGCATAC-----TACTCCGCC 178974
Oy 785 AlaHisGlnMetThrGly----- 790
Db 178973 ACCCAACAAACCACTGCTTCTGTCCTCGAGCAAAACCTTCACGCGTGAATCTCGGGGC 178914
Oy 791 -----GlyIleGlnAlaGlyTYrIleThrPheSerAspAsnPheAsnHisValProTYrArg 808

Accession	Source	Organism	Version	Keywords	Accession	Source	Organism	Version	Keywords
DB 178913	GAAGTCGGGATTTGGCGGGGCTAC-----GGCAGAACCAAGAAATCCCTTTCTT				178863				
QY 809	LeuAlgrPhePheAlaGlyGlyAspGlnSerIleAlaGlyTyrAlaIleHisAspSerLeuSer				828				
DB 178862	GAAGAACCTTCACGGCGGGGCGCTGGGTTCCGGTGGCGGATACGAAGCGGCACGCTCGGT				178803				
QY 829	ProIleSerAspLysGlyTyrIleuThrGlyGlyGlnValLeuAlaValGly-----				845				
DB 178802	CCG-----AAAGTCATATACGAATACAGCGCAAAAAATACGCTACCGCGCGCAACAA				178752				
QY 846	ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp				856				
DB 178751	AAAGCCAAACGTCCTCCGCGAGCTGCTTCCCGATGCCCGCGCAAGACGCGCGCACCC				178692				
QY 857	LeuAlaGlyLeuAlaValPheGlyAspPheIleGlyAsnAlaTyrAsp-----				870				
DB 178691	GTCCGCGTACGCTGTTGGCCGACGAGCAGCGTGTGGACGCAAAACCTACGACGAC				178632				
QY 870	-----				870				
DB 178631	AACACACAGTTCGGCAGCGGCGGAGGTTCAAAACATTTACGGCGCGCAATACCAT				178572				
QY 871	---LysGlyPheThrAsnAspThrIysIleGlyAlaGlyValAlaArgTyrAlaSer				889				
DB 178571	AAATCCACCTTTACACAGCAATTTGGCGTAATCCGCGCGCGCGGTTACCTCGCTCCG				178512				
QY 890	ProValGlyGlnValAlaArgValAspValAlaThrGlyValIleGlySglu-----GluGlyAsn				907				
DB 178511	CCTTTAGGCGCGATGAATTCAGCTACGCGCTACCCGCGTGAAGAAAAACGGAACAGCA				178452				
QY 908	ProIleLysLeuIlePhePheIleGlyThrProPhe				919				
DB 178451	ATCCAAACGCTTCCAAATTCACACTCCGCGCACAGCAGTTC				178416				
RESULT 37									
LOCUS	AF461795								
DEFINITION	AF461795 6444 bp DNA linear BCT 21-JAN-2002								
ACCESSION	Bartonella henselae omp89 (omp89), lpxD (lpxD), fabZ (fabZ), and								
VERSION	lpxA (lpxA) genes, complete cds, and unknown genes.								
KEYWORDS	AF461795.1 GI:18252647								
SOURCE	Bartonella henselae.								
ORGANISM	Bartonella henselae.								
REFERENCE	Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;								
AUTHORS	Bartonella.								
TITLE	1 (bases 1 to 6444)								
JOURNAL	Zimmermann, R., Augustin, K., Schaal, K., and Sander, A.								
REFERENCE	Cloning, nucleotide sequencing, and expression of a hemin-binding								
AUTHORS	protein of Bartonella henselae								
TITLE	Unpublished								
JOURNAL	2 (bases 1 to 6444)								
REFERENCE	Zimmermann, R., Augustin, K., and Sander, A.								
AUTHORS	Direct Submission								
TITLE	Submitted (20-DEC-2011) Institute of Med. Microbiology and Hygiene,								
JOURNAL	Herman-Herder-Str. 11, Freiburg D-79104, Germany								
FEATURES	Location/Qualifiers								
SOURCE	1..6444								
	/organism="Bartonella henselae"								
	/db_xref="taxon:38323"								
	53..1129								
	/note="ORF1"								
	/codon_start=1								
	/transl_table=1								
	/product="unknown"								
	/protein_id="AAL66373.1"								
	/db_xref="GI:18252648"								
	/translation="MIIPIVHEAGHYLIIRWCGIKASVSLGFGPIVGYTDKRGTOIA								
	RLAIIILGGYKFTGDEGLHTGSSQSLPIDVGSFSGSAHAKKAATVPAGFLPVL								

[illegible]

```

VERSION      AF021245.1 GI:2460280
KEYWORDS
SOURCE       Neisseria meningitidis.
ORGANISM     Neisseria meningitidis
              Bacteria; Proteobacteria; beta subdivision; Neisseriaceae;
REFERENCE    1 (bases 1 to 2394)
AUTHORS      Manning,D.S., Reschke,D.K. and Judd,R.C.
TITLE        Omp85 proteins of Neisseria gonorrhoeae and Neisseria meningitidis
              are similar to Haemophilus influenzae D-15-Ag and Pasteurella
              multocida Oma87
JOURNAL      Microb. Pathog. 25 (1), 11-21 (1998)
MEDLINE      98379445
PUBMED       9705245
REFERENCE    2 (bases 1 to 2394)
AUTHORS      Manning,D.S., Reschke,D.K. and Judd,R.C.
TITLE        Direct Submission
JOURNAL      Submitted (27-AUG-1997) Division of Biological Sciences, University
              of Montana, HS 104, Missoula, MT 59812-1002, USA
FEATURES
  source      1..2394
              /organism="Neisseria meningitidis"
              /strain="HH"
              /db_xref="taxon:487"
              1..2394
              /gene="omp85"
              /gene="omp85"
              /codon_start=1
              /product="outer membrane protein Omp85"
              /protein_id="AAC17539.1"
              /db_xref="GI:2460281"
  CDS          1..2394
              /gene="omp85"
              /codon_start=1
              /product="outer membrane protein Omp85"
              /protein_id="AAC17539.1"
              /db_xref="GI:2460281"
  Alignment Scores:
  Pred. No.: 5.47e-07 Length: 2394
  Score: 253.00 Matches: 190
  Percent Similarity: 33.37% Conservative: 121
  Best Local Similarity: 20.39% Mismatches: 343
  Query Match: 5.35% Indels: 278
  DB: 1 Gaps: 40
  US-09-914-168-2 (1-919) x AF021245 (1-2394)
  Oy 116 ThrProLeuSerLeuGluGluPheAlaGlnUserThrGluMetGlyIleAsnPro 135
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
  Db 46 TCGGCTTTGGCATTTGGCCGACTTCCACATCCCAAGACATCCGTCGACGGCTTCACGCGT 105
  Oy 136 AsnAsp-----TyrLeuProGluTyrGlnGlyGluGlnProAsnSer 149
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 106 ACCGAGCGCCAGCAGCATTTACTACTGCGCTCAAGTCGGCGATACCTACAC--- 162
  Oy 150 GluValValValProProThrLeuGlnProGluTyrProGlyLeuIleValSArgLeuTyr 169
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 163 -----GACACACACGCGAGTGCATCATCAACAAAGCGCTGAC 198
  Oy 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186

```

```

  Db 199 GCCACCGTTTCTTGGACGAGCTGCGAATCGCGAGCGGCGAGCTCTGTCACC 258
  Oy 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 259 GTTATCCAGACGCCCCACCACATCGCTCCCTCAATCCATCCAGGCAAAATGCTGCAAAAC 318
  Oy 207 GluProTyrAlaAsnIleLysAlaAlaLeuGluAsp-----IleThrGlnGluSerAla 224
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 319 GAGCCG-----ATTAGAAATAACCTCGAATTCGTTGGCGCTGGGCGAGTCCCAATAC 369
  Oy 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 370 TTTATACGGCGACATCAATCAAGCAGATCGCCGCTGAAAGAAATACCTCGGCGC 429
  Oy 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 430 GCGAAGCTCAATATCCAAATTCACGCCCAAACTGCAACCAACTCGCCGCAACCGGTCGAC 489
  Oy 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 490 ATCGACATACAGATTGACGAGCGCAATTCGCGCAAAATTCACCGACATCGAATTGAA--- 546
  Oy 270 GluGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 547 GGCACCAAGCTATTCGACCGCAAACTGATGCGG----- 582
  Oy 290 LysAlaPheThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 583 -----CAGATGCGCTGACCGAAGCGGCGCATTTGACATGG 618
  Oy 307 -----PheHisGlyLysTyrGluThrLysLysAsnLysLeu 320
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 619 CTGACAGAGAACCAATTCATACGACGACGAAATTTCCCAAGCATGAGAAAGTAACC 678
  Oy 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgThrPheAspArgSerValAspVal 340
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 679 GACTTCTACCAAGAACAGGCTACTGATTCGATTCCTCGATACCGACATCAAAACC 738
  Oy 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlnTyrArg 360
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 739 AACGAGCAAAACCAAGACGACATCAAAATACCGTCGACGAAGGCGGACGTTCCGT 798
  Oy 361 PheAspGluValAlaPhePheThrIleAspProLysThrAsnGlnLeuThrThAspPro 380
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 799 TGGGCGCAAAAGTC-----TCCATCGAAGCGACCAACGAA----- 834
  Oy 381 AspLysLeuProValLysArgGluLeuGlnGluGlnLeuThrValAsnMetGlyGlu 400
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 835 -----GTCCCAAGCGGCACTGGAATAACTGCTGACCATGAAACCCGCCAAA 882
  Oy 401 AlaTyrAsnLeuGlnAlaValAlaIleLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 883 TGGTACGACCAAGCCAGATGACCGCGTTTGGGT----- 918
  Oy 421 AsnMetValAsnThrGluIleValPheProGluArgGlnIleGlnAsnAspGlnVal 440
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 919 -----GAGATTCCAGAACCCGATGCGC 939
  Oy 441 SerPheGlnGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu 460
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 940 TCGGACAGCTACGCAATACG-----GAATTCAGCGTA 972
  Oy 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro--- 479
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 973 CAGCGCGTCCCAAGCGCAAAACCAACCGTCGATTTGCTGCGACATCGAACCGGCGC 1032
  Oy 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492
  Db ::::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
  Db 1033 CGGAATAATCTACGTCGAAGATCCACATCAACCGGCAACAAACCGGACGAA--- 1089
  Oy 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrTrpMetProAspAsp----- 508

```

Db 1090 ---GTGCTGGCGCGCAATTGCGCAATGAAATCGCGCTTACGACACCTCCAAAGCTG 1146
 Oy 509 -----ArgValIleuAlaIleasnHisaspPglValaLsnArgSerIle 523
 Db 1147 CAACGCTCCAAAGAGCGCGCTTGGCTTACTTCCGACACAGACAG----- 1197
 Oy 524 LeuGlyArgIleSeraspAlaValSerAlaValAlaArgAlaIleuProaspGluSer 543
 Db 1198 -----TTTGAATGCCGCTCCCGCTCCGCGACACCCGCAAA--- 1233
 Oy 544 GluAsnGluValIleaspLeuProGluArgThrAlaLeuAlaAsnArgIleThrProAla 563
 Db 1234 -----GTCCATTGTTG-----AACATGAGCCTGACCGACGTTCCACC----- 1269
 Oy 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSeraspLysProArg 583
 Db 1269 ----- 1269
 Oy 584 AspGlyGlnIleGlyLeu-----GlyTyrPglYSeraspThrGlyThrArgLeuValThr 601
 Db 1270 ---GGCTCGCTCGACTTGACGCGCGGCTGGTACAGATACCGGC-----CTGGTCATG 1320
 Oy 602 LysPheGluHisasLeuIleasnArgaspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
 Db 1321 TCCGCAAGCGCTTCCCAAGACAACCTGTCGTACGGGCAAGTCGGCGCGCTCGCGCC 1380
 Oy 622 SerGlaAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
 Db 1381 TCACGAAGCAAA-----ACCAGCGTCACGCGCTGCTGTTTCCGACCGCGTACTTC 1434
 Oy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661
 Db 1435 ACGGCAAGCGGCTGACGCTGGGCTAC-----GATGTTTACGAAAGCAAGCTTCGACCG 1488
 Oy 662 PheaspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly 681
 Db 1489 CGCAAGCAATCGACAGCATCAATATAAACCCACCGCGCA-----GGC 1536
 Oy 682 GlyTyrAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696
 Db 1537 GCAGGATCCGATGACCGCTGTTTACGATACGACCGCGCTGAATTCGGTTGGTG 1596
 Oy 697 LysThrGlnAlaProProGluThrTyrPglAspLeuPro-----ValaspPheVal 713
 Db 1597 GCAGAACACCTGACCTGACACCTTACAAAGCGCGCAACACTATGCCGACTTATC 1656
 Oy 714 Asn-----GlyLys---ProSerGlnGluAlaLeuAlaGlyAlaValaHisLys 730
 Db 1657 AAGCAATACGCAAAACCGACGCGCAGCGAGCTTCAAAAGCTGGCTGTACAAAGCT 1716
 Oy 731 ThrVal-----AlaaspAsnLeuValAsnPrometaArgIleTyrArg 744
 Db 1717 ACCGTGGCGTGGGCGCAACAAACCGACAGCGCTTATGCGCGCGCGCTAC--- 1773
 Oy 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSeraspAlaAsnMetAlaIle 764
 Db 1774 -----CTGACGGCGGTGACGCGCAATC 1797
 Oy 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
 Db 1798 GCCCTCCCGCGAGCAAACTGCAATAC-----TACTCGCGC 1833
 Oy 785 AlaHisGlnMetThrGly-----TACTCGCGC 790
 Db 1834 ACCCAACCAAACTGGTCTTCCCTTAAAGCAAACTTCACGCTGATGCTCGCGCGC 1893
 Oy 791 -----GlyIleGlnAlaGlyTyrIleTyrSeraspAsnPheAsnHisValProTyrArg 808
 Db 1894 GAAGTCGGCATTCGGCGCGCTAC-----GCCAGAACCAAGAAATCCCTTCTTT 1944
 Oy 809 LeuArgPhePheAlaGlyLysAspGlnSerIleArgGlyTyrAlaHisaspSerLeuSer 828
 Db 1945 GAAAACTTACGGCGCGCGCTGGTTCGGTGGGATACGAAAGCGCGACGCTCGGT 2004

Oy 829 ProIleSeraspLysGlyTyrLeuThrGlyGlnValLeuAlaValGly----- 845
 Db 2005 CCG-----AAAGTGTGTGACGAATACGGGAAAAAATACGACTACGGCGCAACAA 2055
 Oy 846 ThrAlaGluTyrAsnTyrGluPheMet-----Lysasp 856
 Db 2056 AAAGCAACGCTTCCCGCGAGCTGCTCTTCCGATGCTGGCGCGAAAGACGGCGGAC 2115
 Oy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
 Db 2116 GTCCCGCTGACGCTTGTGGCAGCAGCAGCGTGGGAGCGCAAAACCTACGACGAC 2175
 Oy 870 ----- 870
 Db 2176 AACACGACTTCCCGCGCGCGCGAGGCTTCAAAACATTTACGCGCGCGCAATACCCAT 2235
 Oy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValAlaArgTyrAlaSer 889
 Db 2236 AAATCCACCTTTACCAAGCAATTCGCTATTCGCGCGCGCGGCTTACTGGCTCTCG 2295
 Oy 890 ProValGlyGlnValArgValaAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
 Db 2296 CTTTATGGCGCGGATGAATTCAGGTACGCTACCGCTGAAGAAAAACCGCAAGACGAA 2355
 Oy 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
 Db 2356 ATCCACGCTTCCCAATTCACACTGGCAGCAGCTTC 2391
 RESULT 39
 AC010535/c 181233 bp DNA 1linear HTG 26-JAN-2000
 LOCUS Homo sapiens chromosome 16 clone RP11-273P15, LOW-PASS SEQUENCE
 DEFINITION
 SAMPLING
 AC010535
 VERSION AC010535.2 GI:6758796
 KEYWORDS HTG; HTGS; PHASE0.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE
 AUTHORS DOE Joint Genome Institute.
 TITLE Sequencing of Human Chromosome 16
 JOURNAL Unpublished
 REFERENCE
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 COMMENT
 Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 On Jan 26, 2000 this sequence version replaced gi:5882403.
 -----Genome Center
 Center: Joint Genome Institute
 Center Code: JGI
 Web site: http://www.jgi.doe.gov

 * NOTE: This record contains 76 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.
 * 1 944: contig of 944 bp in length
 * 945 1570: gap of unknown length
 * 945 1570: contig of 626 bp in length
 * 1571 2381: gap of unknown length
 * 1571 2381: contig of 811 bp in length
 * gap of unknown length

* 2382 2793: contig of 412 bp in length
* gap of unknown length
* 2794 3510: contig of 717 bp in length
* gap of unknown length
* 3511 4621: contig of 1111 bp in length
* gap of unknown length
* 4622 5020: contig of 399 bp in length
* gap of unknown length
* 5021 5969: contig of 949 bp in length
* gap of unknown length
* 5970 7386: contig of 1417 bp in length
* gap of unknown length
* 7387 8141: contig of 755 bp in length
* gap of unknown length
* 8142 9359: contig of 1218 bp in length
* gap of unknown length
* 9360 10926: contig of 1567 bp in length
* gap of unknown length
* 10927 11815: contig of 889 bp in length
* gap of unknown length
* 11816 11979: contig of 164 bp in length
* gap of unknown length
* 11980 14676: contig of 2697 bp in length
* gap of unknown length
* 14677 15050: contig of 374 bp in length
* gap of unknown length
* 15051 16346: contig of 1296 bp in length
* gap of unknown length
* 16347 16482: contig of 136 bp in length
* gap of unknown length
* 16483 17522: contig of 1040 bp in length
* gap of unknown length
* 17523 18676: contig of 1154 bp in length
* gap of unknown length
* 18677 18784: contig of 108 bp in length
* gap of unknown length
* 18785 19464: contig of 680 bp in length
* gap of unknown length
* 19465 20255: contig of 791 bp in length
* gap of unknown length
* 20256 20329: contig of 74 bp in length
* gap of unknown length
* 20330 21427: contig of 1098 bp in length
* gap of unknown length
* 21428 22233: contig of 806 bp in length
* gap of unknown length
* 22234 22614: contig of 381 bp in length
* gap of unknown length
* 22615 22693: contig of 79 bp in length
* gap of unknown length
* 22694 23354: contig of 661 bp in length
* gap of unknown length
* 23355 24389: contig of 1035 bp in length
* gap of unknown length
* 24390 25438: contig of 1049 bp in length
* gap of unknown length
* 25439 26033: contig of 595 bp in length
* gap of unknown length
* 26034 27034: contig of 1001 bp in length
* gap of unknown length
* 27035 28286: contig of 1252 bp in length
* gap of unknown length
* 28287 28664: contig of 378 bp in length
* gap of unknown length
* 28665 29673: contig of 1009 bp in length
* gap of unknown length
* 29674 30889: contig of 1216 bp in length
* gap of unknown length
* 30890 31997: contig of 1108 bp in length
* gap of unknown length
* 31998 33208: contig of 1211 bp in length
* gap of unknown length
* 33209 34186: contig of 978 bp in length

* 34187 36171: gap of unknown length
* contig of 1985 bp in length
* 36172 37717: gap of unknown length
* contig of 1546 bp in length
* 37718 39342: gap of unknown length
* contig of 1625 bp in length
* 39343 41077: gap of unknown length
* contig of 1735 bp in length
* 41078 42930: gap of unknown length
* contig of 1853 bp in length
* 42931 45066: gap of unknown length
* contig of 2136 bp in length
* 45067 46516: gap of unknown length
* contig of 1450 bp in length
* 46517 48277: gap of unknown length
* contig of 1761 bp in length
* 48278 50881: gap of unknown length
* contig of 2604 bp in length
* 50882 52832: gap of unknown length
* contig of 1951 bp in length
* 52833 54941: gap of unknown length
* contig of 2109 bp in length
* 54942 56865: gap of unknown length
* contig of 1924 bp in length
* 56866 59370: gap of unknown length
* contig of 2505 bp in length
* 59371 62658: gap of unknown length
* contig of 3288 bp in length
* 62659 64918: gap of unknown length
* contig of 2260 bp in length
* 64919 67208: gap of unknown length
* contig of 2290 bp in length
* 67209 69548: gap of unknown length
* contig of 2340 bp in length
* 69549 72168: gap of unknown length
* contig of 2620 bp in length
* 72169 74419: gap of unknown length
* contig of 2251 bp in length
* 74420 78177: gap of unknown length
* contig of 3758 bp in length
* 78178 81052: gap of unknown length
* contig of 2875 bp in length
* 81053 85165: gap of unknown length
* contig of 4113 bp in length
* 85166 90874: gap of unknown length
* contig of 5709 bp in length
* 90875 94813: gap of unknown length
* contig of 3939 bp in length
* 94814 98166: gap of unknown length
* contig of 3353 bp in length
* 98167 102653: gap of unknown length
* contig of 4487 bp in length
* 102654 108499: gap of unknown length
* contig of 5846 bp in length
* 108500 113109: gap of unknown length
* contig of 4610 bp in length
* 113110 118023: gap of unknown length
* contig of 4914 bp in length
* 118024 126141: gap of unknown length
* contig of 8118 bp in length
* 126142 134123: gap of unknown length
* contig of 7982 bp in length
* 134124 141203: gap of unknown length
* contig of 7080 bp in length
* 141204 149918: gap of unknown length
* contig of 8715 bp in length
* 149919 158297: gap of unknown length
* contig of 8379 bp in length
* 158298 167935: gap of unknown length
* contig of 9638 bp in length
* 167936 181233: gap of unknown length
* contig of 13298 bp in length.
! Location/Qualifiers

FEATURES

```

source
1. 181233
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="16"
   /clone="RP11-273P15"

BASE COUNT    45888 a 45036 c 44713 g 45408 t    188 others
ORIGIN

Alignment Scores:
Pred. No.:      0.00049      Length:      181233
Score:          245.50       Matches:      130
Percent Similarity: 37.00%   Conservative: 55
Best Local Similarity: 26.00% Mismatches:      180
Query Match:     5.19%      Indels:        136
DB:              2          Gaps:           21

US-09-914-168-2 (1-919) x AC010535 (1-181233)

OY 501 H1sLeuTYrAspMetProAspArgValLeuAlaIleAsnHisAspAsp----- 517
    |||||
Db 26074 CATCTTTACAGAGGAGGAGCTGATAGCATCTGACACAAAAGCCATCATATTTCGGGCC 26015
OY 518 -----GlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAla 535
    |||||
Db 26014 ACATGCTACCCATCTGTGGCACTGCTGGAGATTCTGCAGTGCACATGCACCTTGC 25955
OY 536 Arg-----AlaIleLeuProAsp 541
    |||||
Db 25954 AAATGATTCATTCACAGAAAGTCACAGTGCAGTGCATCTACACACATTTCTCCACAG 25895
OY 542 GluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThr 561
    |||||
Db 25894 GAGCAG-----CATCTCTCTCATTGACCAAGATTCCTTCCAAACATCAC 25850
OY 562 ProAlaAspValTyrGlnSer-----LysLysValProLeuTyrValPheValAlaSer 579
    |||||
Db 25849 AATATTAAGCAATTTGGGAGAACTGTCAGAGGATTTCTTTCAGCTGTATTCAGAGTCC 25790
OY 580 AspLysProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeu 599
Db 25789 CAAATTACTAGCAGCAGACATGGCCCACTGGTGTGG-----CAGACCATGAGACAGCGC 25736
OY 600 ValThrLysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlyLeu 619
    |||||
Db 25735 -----CCACCGGGCTATTTCGCG----- 25718
OY 620 ArgLeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisPro 639
    |||||
Db 25717 -----AGCCAGCCCGAGGTTGCGCAGTACAGAAACATCTGCTTC----- 25679
OY 640 LeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThr 659
    |||||
Db 25678 -----CTTCCTTCAACACTGGGCAAGAGAGAGAAC-----CATGCTCTC 25637
OY 660 AsnGlyPheAspLeuSerThrArgThrLeuGlnHisGlyIleSer----- 674
    |||||
Db 25636 AATGAATACTCACTGGCAGTACATTCTGAGACACCCATGAGTACTATTCTTCT 25577
OY 674 ----- 674
Db 25576 CTTCGACAGCTACTTTTACAAACATTCCTTCCTCCATTTAGAGCTGCTGGGATTTTA 25517
OY 675 -----ArgSerIleIle-----GlnAsnGlyGly 682
    |||||
Db 25516 AAGCTTCAGAAACCCAGAGCTTTATTAATAAGAAAATGTAACCTGGCAGAAATTTTGGG 25457
OY 683 TrpAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProPro 702
    |||||
Db 25456 -----GTGTGACACCGTAATTCCTTCCTGCTGTCCTGTCCTGCTGCTGCTTCT 25406
OY 703 GluThrTrpGlnAspLeuProValAspPheValAsnGlyLysProSerGlnIleValLeu 722
    |||||
Db 25405 GTCCCGTGG-----CTTCGCCCGGTGTTGTTGAGTGAAGAGAGAAAGAGG----- 25358

```

```

OY 723 LeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGly 742
    |||||
Db 25357 -----ACAGAAAAAACTTAACGACATGCGCGGTGATTAACCGGATTTTTCG 25311
OY 743 TyrArgGlnArgTyrSerLeuGlnValGlySerSerGlyLeuValSerAspAlaAsn-Me 762
    |||||
Db 25310 TTAGCCGCGCGC-----TCCCTTGTTCGCGTTCT-----ATGAGTGCAGCTGATATTC 25260
OY 762 ValAlaIleAlaArgAlaGly-----IleSerGlyValTyrSerPheGlyAspAsnAlaTyr-- 780
    |||||
Db 25259 GGGGAGGTAGCGCCCTCTTTGTTGTGGGGCATACACGTTGGGTCACCAACTATTTGA 25200
OY 781 -----GlySerAsnArgAlaHisGlnMetThrGlyGlyIleGln----- 793
    |||||
Db 25199 GACACCCACGGGGGGTTAAACCGATTTCCCGTTTACAGGCGCAGAACGCTGATTC 25140
OY 794 -----Ala-GlyTyrIleTyrSerA 800
    |||||
Db 25139 GCACACTGTACGATCGCATGTTTGTTTACACGGCGACGCGTGGCTGATTTGAACCG 25080
OY 800 sPAsnPhAsnHisValProTyrArgLeuArgPhePheAlaGlyLysPcInserTlea 820
    |||||
Db 25079 GTGATTTGACAAAGTACCGCGCGGATCTGCTTCTTCGCGCGGGCGCAGCCACTATTTC 25020
OY 820 rGGLYTYrAlAHIsAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGly 840
    |||||
Db 25019 GTGCTTAATATCAATCTATCTGCTCCGAATACGCCAACCGTACCTGAAAGGGGCTT 24960
OY 840 InValLeuAlaValAlaGlyThrAlaGlyTyrAsnTyrGluPheMetLysAspLeuAlaGly 860
    |||||
Db 24959 CGAATGTATTAACCGGATCGGTGAATACAGTACACAGTGCACCGAAATGCTGGGCG 24900
OY 860 lAlaValPheGlyAspIleGlyAsnAlaTyrAspGlyPheThrAsnSprThrLysIleG 880
    |||||
Db 24899 CGGTGTTGTGCATGTGCGCAGGCGGTGAACGATATTCGCCGCGCAGCATTTAAACCG 24840
OY 880 lYAlaGlyValAlaGlyValArgTyrAlaSerProValGlyInValArgValAspValAlaT 900
    |||||
Db 24839 GTACCGGGGTGCGCGCTGCGCTGGAATGCGCGCTGCGGCGCAATCAATCATTTGGCG 24780
OY 900 hrgLysVal-----LysGluGluGlyAsnProIleLysLeuHisPhePheIleGly 916
    |||||
Db 24779 TACCGGTGCGGATTAAGACGACACGCG-----TTACAGTTTATCAATCGCT 24733
RESUT 40
AE012236/c
LOCUS      AE012236      10029 bp      DNA      linear      BCT 23-MAY-2002
DEFINITION Xanthomonas campestris pv. campestris str. ATCC 33913, section 144
of 460 of the complete genome.
ACCESSION  AE012236 AE008922
VERSION    AE012236.1 GI:21112424
KEYWORDS
SOURCE
ORGANISM
Xanthomonas campestris pv. campestris str. ATCC 33913.
Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
Xanthomonas.
1 (bases 1 to 10029)
da Silva,A.C.R., Ferro,J.A., Reinach,F.C., Farah,C.S., Furlan,L.R.,
Quaggio,R.B., Monteiro-Vitorello,C.B., Van Sluys,M.A., Almeida
Jr.,R.F., Alves,L.M.C., do Amaral,A.M., Bertolini,M.C.,
Camargo,L.E.A., Camarotte,G., Camarvan,F., Cardozo,J.,
Chambergro,F., Clapina,L.P., Cicarelli,R.M.B., Coutinho,L.L.,
Cursino-Santos,J.R., El-Dorriy,H., Faria,J.B., Ferreira,A.J.S.,
Ferreira,R.C.C., Gruber,A., Katsuyama,A.M., Kishi,L.T., Leite
Jr.,R.P., Lemos,E.G.M., Lemos,M.V.F., Locali,E.C., Machado,M.A.,
Madeira,A.M.B.N., Martinez-Rossi,N.M., Martins,E.C., Meidanis,J.,
Menck,C.F.M., Miyaki,C.Y., Moon,D.H., Moreira,L.M., Novo,M.T.M.,
Okura,Y.K., Oliveira,M.C., Oliveira,V.R., Pereira,Jr.,H.A.,
Rossi,A., Sena,J.A.D., Silva,C., de Souza,R.F., Spinola,L.A.F.,
Takita,M.A., Tamura,R.E., Teixeira,E.C., Tezza,R.I.D., Trindade dos
Santos,M., Trufl,D., Tsai,S.M., White,F.F., Setubal,J.C. and

```

TITLE	Kitaajima, J.P.
	Comparison of the genomes of two <i>Xanthomonas</i> pathogens with differing host specificities
JOURNAL	Nature 417 (6887), 459-463 (2002)
MEDLINE	22022145
PUBMED	12024217
REFERENCE	2 (bases 1 to 10029)
AUTHORS	da Silva, A.C.R., Ferro, J.A., Reinach, F.C., Farah, C.S., Furlan, L.R., Quaggio, R.B., Monteiro-Vitorello, C.B., Van Sluys, M.A., Almeida Jr., N.F., Alves, L.M.C., do Amaral, A.M., Bertolini, M.C., Chambergo, L.E.A., Camarotte, G., Cannavan, F., Cardoso, J., Chambergo, F., Chapina, L.P., Cicarelli, R.M.B., Coutinho, L.L., Cursino-Santos, J.R., El-Dorri, H., Faria, J.B., Ferreira, A.J.S., Ferreira, R.C.C., Ferro, M.I.T., Formighieri, E.F., Franco, M.C., Greggio, C.C., Gruber, A., Katsuyama, A.M., Kishi, L.T., Leite Jr., R.P., Lemos, E.G.M., Lemos, M.V.F., Locati, E.C., Machado, M.A., Medeira, A.M.B.N., Martinez-Rossi, N.M., Martins, E.C., Meidanis, J., Menck, C.F.M., Miyaki, C.Y., Moon, D.H., Moreira, L.M., Novo, M.T.M., Okura, V.R., Oliveira, M.C., Oliveira, V.R., Pereira Jr., H.A., Rossi, A., Sena, J.A.D., Silva, C., de Souza, R.F., Spínola, L.A.F., Takita, M.A., Tamura, R.E., Teixeira, E.C., Teze, R.I.D., Trindade dos Santos, M., Truffi, D., Tsai, S.M., White, F.F., Setubal, J.C. and Kitaajima, J.P.
TITLE	Direct Submission
JOURNAL	Submitted (28-NOV-2001) Departamento de Biotecnica, Universidade de Sao Paulo, Av. Prof. Lineu Prestes 748, Sao Paulo, SP 05508-900, Brazil
FEATURES	
source	1. 10029 /organism="Xanthomonas campestris pv. campestris str. ATCC 33913" /strain="ATCC 33913" /db_xref="ATCC:33913" /db_xref="taxon:190485" /note="pathovar: campestris" complement(72..2528) /gene="oma" /note="XCC1365" complement(72..2528) /gene="oma" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark" /codon_start=1 /transl_table=11 /product="outer membrane antigen" /protein_id="AA040663.1" /db_xref="GI:21112425" /translation="MKLLHSFRLSSSYAGTQOHPSTPGCDMTRLPTRRLALALAAGISLPAALAAEPVADSIKVDGLRIAGSGVETLTPVNRGDTVDGCGVDSIRLKYRGFEDVOLDRQGNILVTYKRAIPINKLITVGNKDISEELKGLGIGLIEGTFDRSLDRVTQELRTOINNRKATNVEITPTVSPUDRNRVATAIKGKAKIRHVLIGTEKFNEDIILENWESEKHNMAWYRQDOYSEKMSGDLKLNWYLDRGVDFSVDSIOVISPKRDMFVAGVTEGOYKISEIKYGTGDTLPADIERLVIKADIDISRALLEYSDAITNTLSNYSYAFKAVNPPIPTNREDRIVANLOVGPVSVRRLLKVGNSRSDVLRERMKROFENTWYSOAATIDRSKIRLORLGYSPSVDETPAVGSDOYDVYVVKETSSGFVFGLCFSQAFCTTTSVQLSQNNPLGGGRVAVEASRTFQQRKAFSYNPFPTDGVSLGYNLSKRELDYSDNTAOFNSTNGSAOVVGVGPTMDTVSAMGVDSNQITTFPGTTPQALVDYIOAIGDTFAVTRTEGMAADTRNDPFMPRNGVORGLLETTLPSTVEYKLANOISNWPPIIPALVNTREFELGDSYGDSSGVTIADGTTRPVTATGLPEYENYAGTNSVGFEDNTLGPSEALINGRNGOPLGGSLKTVGSVMYPKLIDEPSPARISAFDPFGNYSVDVAFKANELRASTGVALLMRAPVGVISISYAFPLKKEDNEIERLQFTFCGF"
gene	complement(2525..3871) /gene="XCC1366" complement(2525..3871) /gene="XCC1366" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
CDS	/codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AA040664.1" /db_xref="GI:21112426"
gene	/translation="MGDFIGSVMMIVTSLGVLTVTFHEFGHFVARRCGVKVLFVSFVGKPLMRRDRHRTGEFAIAAIPLOGYVKMLDERGEVHPAREQAFNRKTVIMORIAVAACPIANILLCMVLMAMFVIGKODYSATGRADGLAAAGLVGSEITVLDGYSVMSDASMTLTAAMDRDVOYLTAAREGSEHPTIRLSOLPAGDERVATLADGQMGFLOPPVYDKVYGAADGVLCPRDRIYADIGOPFRSAGEVPAQLOALGTGGGMTEVAREDDRILALEIAPRKSPGGMGLGVFAATAAPYDVSROOYGAFAAPAARIREGKMTADSLGMMKRMILTGQASVKNISGPVTIARAANASERVDWFLVFLGLSLSLIITNMPILPDGHLIYVLEIELVKSIPSERAMIGOVGLAVLGLMGLAFVNDILGLVPR" complement(3898..5088) /gene="dxr" /note="XCC1367" complement(3898..5088) /gene="dxr" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
CDS	/codon_start=1 /transl_table=11 /product="1-deoxy-D-xylulose 5-phosphate reductoisomerase" /protein_id="AA040665.1" /db_xref="GI:21112427" /translation="MSGSLRRVAVFGANGSIGASALVDYIARHPERLRASVLSGSKVDALLALCVLHRPAHVAYADAAIYPAALDGLRAAGLTTOAHAGDQALDAAASDAPCTVVAIVGAGLSSTLAANAAGKRLILANKRESIYAGELLTRTAANAAGATITIPDSHSAIFQCLASCDASRCVRRVITLTAAGCPFRGROALAVTTPQAQVAHPKMSGKRISYDSATLMNGLVETIENHLFGLPGEIDVLVHPQSLVSLVEFVGSTLAQGLDPMRTTLAVGLAMPREVESGVGIDLQGRIDFEAPDEAFPCRLADALDARAGCTAAVILMAANEVAVSAPFLQGVKGFALIPALVEHTLTTRQADRTLNLALFDADBARRTTBALAHSLHA"
gene	complement(5091..5918) /gene="cdsa" /note="XCC1368" complement(5091..5918) /gene="cdsa" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
CDS	/codon_start=1 /transl_table=11 /product="phosphatidate cytidyllyltransferase" /protein_id="AA040666.1" /db_xref="GI:21112428" /translation="MTQFNVIALIIMAPLCAICVLLPTOMIALAALILFLTGLWEMLKLSGIDDSLPRTVILNLINLWIMVWASGVSVFQIATVAGVGMVLAALIMLRFNFGADHNGCARAKLKAAGTLVYPRMAALVLIHANPDKNMILTLTRVYWAADSAVFAGRRGKHLAPRISPNKTTISLGLGLVAGIACAFGMGLTLAHVPQVLVAAYAVLASVLGDLFESLILKRNHAGAKSGVTIVGHGVLDRIDGVLAALPVFALGKEIIGF"
gene	complement(5915..6691) /gene="upps" /note="XCC1369" complement(5915..6691) /gene="upps" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
CDS	/codon_start=1 /transl_table=11 /product="undecaprenyl pyrophosphate synthetase" /protein_id="AA040667.1" /db_xref="GI:21112429" /translation="MASTAMHSEPSAAVPRHLAIIIMDNGRMAORRRRRPVYIGHRAGAAVAVITIEFCLEKGVSAITLFAFSSENMGRQDEVDALMKLFILHALDGVDELORGVQVVRFGDSRFEAPLRDMAGERTTANTRTVLSIASYGGRODIAWAAALAVEYAGRLROPEDIDENALLASRYALADIPARPDLFIRGQGTIRSNGLMDLATTELMFTETLWPEFDAGVLQDALDDYTAGRERRGTLTSAOIAEKATEASSA"
gene	complement(6694..7251) /gene="ftr" /note="XCC1370" complement(6694..7251) /gene="ftr" /note="identified by sequence similarity; putative; ORF located using Blastx/Glimmer/Genemark"
CDS	/codon_start=1 /transl_table=11 /product="ribosome recycling factor" /protein_id="AA040668.1"


```
OY 677 IleIleGlnAsnGlyGly-----TyrPasnArgThr 686
      |||||
Db 908 AACAGCACGAATGGGTGGCGCACAGTGTGTCGGCGTCCGCGATCACCAGACAGATACT 849
OY 687 TyrSerLeuArgTyrArgLeuAspLysIleuLysThrGlnAlaProGluThrTrpGln 706
      ||| | |
      |||||
      |||||
Db 848 GTCTCGGCGATGATCGGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 789
OY 707 AspleuProValAspPheValAsn-----GlyLysProSerGlnGluAlaLeuLeuAla 724
      |||||
      |||||
      |||||
Db 788 CAGGCAATTGTCGACTACATTCAAGCATGCGACCGCATCTTTCATGAGCGCGGCGACT 729
OY 725 GlyValAlaValAlaLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
      |||||
      |||||
      |||||
Db 728 GAGTTCGGTGGCGCGCGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
OY 745 GlnArgTyrSerLeuGluVal-----GlySerSer----- 754
      |||||
      |||||
      |||||
Db 668 CACGCTGTGGCGCTGGAACACCGCTTCTGCTTCCACGCGTGAAGTACTACACGCTTAAT 609
OY 755 -----GlyLeuValSerAspAlaAsnMetAla 763
      |||||
      |||||
      |||||
Db 608 TATCAGATTTCGAACACTGCGCGCATCTTCTGCTTCTCAATACCGCGTTTGAG 549
OY 764 IleAlaArgAlaGlyIleSerGlyValTyrSerPheGly-----AspAsnAlaTyrGly 781
      |||||
      |||||
      |||||
Db 548 TTGCGCTATGCGCGACACTGACGACGACGACGACGACGACGACGACGACGACGACGAC 489
OY 782 SerAsnArgAlaIleGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAspAsn 801
      |||||
      |||||
      |||||
Db 488 ACCAGCGCTCGGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 462
OY 802 PheAsnIleValProTyrArgLeuArgPhePheAlaGlyIleAspGlnSerIleArgGly 821
      |||||
      |||||
      |||||
Db 461 -----CTGCCCTTCTACGAGAACTTCTACGCTGCGCGTACCAATTCGTCGCGCGC 411
OY 822 TyrAlaIleAspSerLeuSerProIleSerAsp-----LysGlyTyr 835
      |||||
      |||||
      |||||
Db 410 TTCGAGACGACACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 351
OY 836 LeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLys 855
      |||||
      |||||
      |||||
Db 350 CCACCTCGCGCGCTGCTGAAACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 291
OY 856 Asp-----LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
      |||||
      |||||
      |||||
Db 290 GACAGCCCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 231
OY 870 AspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer 889
      |||||
      |||||
      |||||
Db 230 GATCTTTCAAGGCCCAATGACCTGCTGCCCAACCGCGTGTGCGCATCTGTCGCGCGCGCG 171
OY 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlnGluGlyAsnPro--- 908
      |||||
      |||||
      |||||
Db 170 CCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 111
OY 909 ---IleLysLeuHisPhePheIleGlyThrProPhe 919
      |||||
      |||||
      |||||
Db 110 ATCGAGCGCTGTGCACTTACCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 75
```

Search completed: May 6, 2003, 01:32:46
Job time : 5790 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OW protein - nucleic search, using frame_plus_p2n model

Run on: May 5, 2003, 23:48:35 ; Search time 325 Seconds

(without alignments)
6367.958 Million cell updates/sec

Title: US-09-914-168-2

Perfect score: 4727
Sequence: 1 MSKPVLFANRSEMPVALAAY.....TGKVEGPNIKLHFICTPF 919

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+p2n.model -DEV=x1h
-Q=/cgr2.1/USFTO.spool/US09914168/runat_28042003_151614_6775/app_query.fasta.1.1095
-DB=N.Geneseq.101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=ISO9914168 @CGN.1.1.205 @runat_28042003_151614_6775 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEOUTR -NEG_SCORES=0 -MAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-Ygapop=10 -Ygapext=0.5 -Delop=6 -DELEXT=7

Database :

N.Geneseq.101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4727	100.0	2760	21	AAA50536	Moraxella catarrha
2	4727	100.0	89047	22	AAF28547	Genomic fragment #
3	4581	96.9	2670	21	AAA50537	Moraxella catarrha
4	604.5	12.8	1374	22	AAF61044	P. putida KT2440-a
5	436.5	9.2	1737	21	AAA50269	Haemophilus influe
6	427.5	9.0	1731	21	AAA50270	Haemophilus influe
7	320	6.8	2193	23	AAAS93089	DNA encoding novel
8	319	6.7	944	24	ABK5155	Haemophilus influe
9	265.5	5.6	2379	22	AAAS07278	Neisseria meningit
10	265.5	5.6	2379	22	AAH42129	Neisseria meningit
11	259.5	5.5	2505	21	AAAI1535	DNA encoding sequenc
12	259	5.5	2394	22	AAAS07277	Neisseria meningit
13	259	5.5	2394	22	AAAS07279	Neisseria gonorrhoe
14	259	5.5	2394	22	AAH42128	Nucleotide sequenc
15	259	5.5	2394	22	AAH42130	Nucleotide sequenc
16	259	5.5	53253	21	AAAB1478	N. meningitidis pa
17	259	5.5	349980	21	AAEF21544	Neisseria meningit
18	253	5.4	2394	21	AAAI15156	DNA encoding outer
19	238.5	5.0	6617	22	AAI97964	Lawsonia intracell
20	235.5	5.0	1425	23	AAAS81897	DNA encoding novel
21	235.5	5.0	1425	23	AAAS93087	DNA encoding novel
22	229	4.8	2325	20	AAAX91724	Porphyromonas ging
23	229	4.8	2409	20	AAAX91597	Porphyromonas ging
24	224.5	4.7	1830	21	AAAB45507	Neisseria meningit
25	224.5	4.7	1830	21	AAAB45508	Neisseria meningit
26	222.5	4.7	1848	21	AAZ53624	Neisseria meningit
27	216.5	4.6	1764	21	AAAB45509	Neisseria meningit
28	216.5	4.6	1848	21	AAZ53623	Neisseria meningit
29	216.5	4.6	92934	21	AAAB1473	N. meningitidis pa
30	216.5	4.6	173325	21	AAAF21613	Neisseria meningit
31	214.5	4.5	2388	24	ABO72980	Methylococcus caps
32	214.5	4.5	2351	22	AAO90024	H. capsulatus
33	206.5	4.4	2751	22	AAZ25588	H. pylori Hps120 e
34	205.5	4.3	3012	19	AAV52012	Helicobacter polyp
35	200	4.2	2778	18	AAZ5138	H. pylori cell env
36	200	4.2	2778	20	AAAX75802	H. pylori outer me
37	184.5	3.9	2989	15	AAO66202	H. influenzae PAK
38	182	3.9	2342	23	AAAS93745	DNA encoding novel
39	181	3.8	1916	23	AAAS93088	DNA encoding novel
40	178.5	3.8	2974	15	AAO65201	H. influenzae SB33
41	176.5	3.7	2442	21	AAZ28550	M. catarrhalis (AT
42	176.5	3.7	2442	21	AAZ28551	M. catarrhalis (AT
43	176.5	3.7	99629	22	AAZ28550	Genomic fragment #
44	174.5	3.7	2019	22	AAAF61065	P. putida KT2440-a
45	173.5	3.7	8922	23	ABL04053	Drosophila melanog

ALIGNMENTS

RESULT 1
AAA50536
AAA50536 standard: DNA: 2760 BP.

05-DEC-2000 (first entry)

Moraxella catarrhalis BASB081 gene coding region.

BASB081 gene; infection; otitis media; sinusitis;
inflammation; therapy; antibacterial; antiinflammatory; vaccine;
diagnosis; ds.

Moraxella catarrhalis.

Key Location/Qualifiers
CDS 1..2670

Db	121	CAATGACCCGCTCATGATGACACCGCCATCATCAACAAAGCAGGCAATCCGCTGTTTG	180
Qy	61	LeuThrProGluGlnIleGlnAlaArgLeuAsnAlaIleGluLeuAsnAlaLysProGln	80
Db	181	CTAACACCGCTGAGAGATACAAAGCAACGCGCTTATCTCTGACTGATGATGAAGCCCA	240
Qy	81	SerGlnAlaLeuAspValValAsnPheAspGlnSerProIleSerArgIleGlyIlu	100
Db	241	TCACAGCGCTTGGATGGTGTCAATTTTGATGATCAATCGCCGATCTGTAATCGGTGAG	300
Qy	101	GlnSerProProLeuGlyLeuAspMetSerValIleGlnGlnThrProLeuSerLeu	120
Db	301	CAATACACCCCTTTGGTGGTTTGATGTGCTGCTCAACAAACACACCGGTAGCCTTG	360
Qy	121	GluGluLeuPheAlaGlnGlnSerThrGluMetGlyIleAsnProAsnAspTyrIlePro	140
Db	361	GAGGAAATTAATTTGGCTCAAGAAATCTACTAGATGGGAAATCAATCCAATGATTATATCCA	420
Qy	141	GluTyrGlnGlyGlnGlnProAsnSerGluValValProProThrLeuGluProGlu	160
Db	421	CAATTCACAGCGAGGACCAACCTTAATAGAGCTGGTGTACCCGACATTAAACCTGAA	480
Qy	161	LysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspValAlaAsnLysVal	180
Db	481	AAACAGAGGTGTGATCAACAGCGCTTATGACGCGCTATTATATGATGTTGTCATTAAGTG	540
Qy	181	ProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyIluThrSerAlaIleGly	200
Db	541	CCTAGCGTTAAAGGCAAAATTTATCAATCATCCAAATCAGGCCAAACAGTCGATGGG	600
Qy	201	SerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaIleLeuGluAspIleThr	220
Db	601	TCATCGCATCAAAAAGAGAGCCCTTATGCAAAATCAAGACAGCACTTAAGACATCAAC	660
Qy	221	GlnGlnSerAlaMetLaspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuVal	240
Db	661	CAAGATCAGCGATGATGATTTGAATGGCTCATCCACGCTTAAGCAAACTCTTTGGTG	720
Qy	241	AlaAlaArgAlaValAlaGlyTyrTyrAspIleAspLeuSerIleIleArgSerIleGly	260
Db	721	GCACGCGGTGCTGCTGCTATTAATGATATGATTTATCAATCAATAAAGATGATCGGA	780
Qy	261	GluValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaVal	280
Db	781	GAGGTGATGTTCATCAATCCATGATTAAGGGAACCGTTTATATGATTTGACGCGGTG	840
Qy	281	GluValArgIleGlyGlnGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValPro	300
Db	841	GAGTACGAGGTGAAGTGTGATGATTAAGCATTTACTACCGTGGCGATGAGGTGCCA	900
Qy	301	LeuLeuIleGlyAspValPheHisIleGlyLysGlyLysThrLysLysAsnLeuIleGlu	320
Db	901	TTTGCTGATCGGCATGCTTTTCATCTATGCGCAGTACGAACCAAAAAAATCTCATGCA	960
Qy	321	AsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgTrpLeuAspArgSerValAspVal	340
Db	961	AATGCCAGTCTGATACATGATATTTTCATGGCGGTGGCTGGATGCTTCAATGATGATA	1020
Qy	341	IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArg	360
Db	1021	ATTTTTCGCAATTAATCCGCTGATGTCACCTTAATTTATGATACAGTACGAGATGCC	1080
Qy	361	PheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro	380
Db	1081	TTTGATGAGGTGATTTTAAACATTAAGTAAACCAATTAATTAACACCATGCA	1140
Qy	381	AspLysLeuProValLysArgIleLeuLeuGlnGlnLeuLeuThrValAspMetGlyGlu	400
Db	1141	GATTAAGTGGCAATTAAAGAGATTAATTAATGATTTGATTTGATTTGATTTGATTTG	1200
Qy	401	AlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe	420
Db	1201	GCTTACATTTACAGCGGTGGTGCATCTTAATTAATGATTTGATTTGATTTGATTTG	1260

QY 421 AsnMetValAsnThrGluIleValPheProGluIuArgGluGlnIleGlnAsnAspGluVal 440
 |||||
 Db 1261 AATATGGTGAATACCGAGATGTCTTCCAGAGGGTGAACAGATCCAAAGACCAACTGTG 1320
 QY 441 SerPheGluInSerSerSerSerArgThrGluProAlaGlnValAspGluSerThrIeu 460
 |||||
 Db 1321 AGCTTTAGAGCATCTTCAAGTAGCCGTAACACAGCAGCAAGTTGATGAAGACACTT 1380
 QY 461 GluProValIleGluThrValGluIleThrAspGlyIleLeuMetAspIleSerProIle 480
 |||||
 Db 1381 GAACCTGTCAATGAAACCGTTGAGCTTAACGAGATGGGATTTAAATGATATTTCCGCCATC 1440
 QY 481 GluPheSerAlaSerAsnIleuIleGlnAspLysIleuAsnIleuValAlaAlaLysAlaArg 500
 |||||
 Db 1441 GAATTTAGTGATCTAATCTGATTCAGACAGCAAGTAATTTGGTGGCTCCCAAGGCTCGC 1500
 QY 501 HisLeuTyrAspMetProAspAspArgValIleuAlaIleAsnHisAspAspGlyValAsn 520
 |||||
 Db 1501 CATTTAATATGACATGCCATGATAGGGTCTGCTGCATCAATCATGATGATGGCTAAT 1560
 QY 521 ArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleuPro 540
 |||||
 Db 1561 CGCTCTATTTGGCGACAGATACGCGATGCCATCTGCCGTTGACACGCTGCTATTTACT 1620
 QY 541 AspGluSerGluAsnGluValIleAspLeuProGluIuArgThrAlaIleuAlaAsnArgLys 560
 |||||
 Db 1621 GATGAATCTGAATAATGAGGTATAGATTGGCCGACGCTACCGCATGGCTAATGCCAAG 1680
 QY 561 ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAsp 580
 |||||
 Db 1681 ACCCTCTGCTATGCTATCAAAAGTAAAAAGTGGCGCTATGTCTTGTGGCGAGTAT 1740
 QY 581 LysProArgAspGlyGlnIleGlyLeuGlyTyrPglSerAspThrGlyThrArgLeuVal 600
 |||||
 Db 1741 AAACACAGAGATGTCAAATGGTTGGCTGGGAGATGGACACAGGTAAACCGCTACTTC 1800
 QY 601 ThrLysPheGluHisAsnIleuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArg 620
 |||||
 Db 1801 ACAAAATTTGAGCATAAATTTGATTAATCGTATGCGTATCAAGCAGCGCTGAGCTAAGA 1860
 QY 621 LeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeu 640
 |||||
 Db 1861 CTGTCTAGAGTAAATAAGGGGTCAAGTATATGCCACAAACCGCTTAGCCACCTCTTA 1920
 QY 641 AsnAspGluLeuArgAlaThrLeuGlyTyrGlnGlnIleuValPheGlyHisSerThrAsn 660
 |||||
 Db 1921 AATGATCAGCTAAAGCAACTTTGGGTTATCAACAAGAGTTTGGTCACTTACCAAT 1980
 QY 661 GlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsn 680
 |||||
 Db 1981 GGTTTGATTTATCCACACGACCCTAGCATGATTAACCCGAGATTAATCCAAAT 2040
 QY 681 GlyLysThrAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLysLeuThrGlnAla 700
 |||||
 Db 2041 GGTCGCTGGAAATCTACTTATTCATTCGTGATTCGTCTTGATGAAGCTTAAAAACCAACA 2100
 QY 701 ProProGluThrTyrPglAspLeuProValAspPheValAsnGlyLysProSerGlnIle 720
 |||||
 Db 2101 CCCCCTGAACATGGCAGAGATTTACCACTGATTTTGCATATGTTAGCCCAAGCCAAAGAG 2160
 QY 721 AlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnIleuValAsnProMet 740
 |||||
 Db 2161 GCGTTATTGGCAGGTGTGCTGCATTAACGTTGAGAGATTAATTTGTTAATCCGATG 2220
 QY 741 ArgGlyLysArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAla 760
 |||||
 Db 2221 CGTGGCATATGTCAGCATATCTTTAGAGGTTGGCTCAACCGGTTTGGTTCGATCT 2280
 QY 761 AsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyr 780
 |||||
 Db 2281 AATATGGCTATTTGCTCGAGCTGGTATTAAGTGGCGTATTAAGTTTGGGGGATTAAGCTTAT 2340

QY 791 GlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleThrSerAsp 800
 |||||
 Db 2341 GGCACCAATTCGCGCCATCAGATGACTGTGGCATACAAAGCAGATTAATTTGGTGGAT 2400
 QY 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyValAspGlnSerIleArg 820
 |||||
 Db 2401 AATTTTAATCATGTCCATATCCGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2460
 QY 821 GlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGln 840
 |||||
 Db 2461 GCATATGCACATGACAGATTTATACCTATATCAGATTAAGGTTATCTGACAGCCGCTCA 2520
 QY 841 ValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAla 860
 |||||
 Db 2521 GATTTGGCGGTTGGTACAGCTCAATATATATGATTAATTAAGATTTGGCTTTGGCG 2580
 QY 861 ValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGly 880
 |||||
 Db 2581 GTTTTGGTGTATTTGGTAAATGCTTATGATTAAGGCTTTACTAATGAATACCAAAATTTGGT 2640
 QY 881 AlaGlyValGlyValArgTyrPheLysProValGlyGlnValArgValAspValAlaThr 900
 |||||
 Db 2641 GCAGGTGCGGTGCTTCCGCTGGCATCACCTGTGCGTCAAGTTGCTGTTGATGTGGCACT 2700
 QY 901 GlyValLysGluGluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
 |||||
 Db 2701 GGTGTCAAGAAAGAGGCAATCCATTAAGCTGATTTTATTTGGCACACCACTT 2757
 RESULT 2
 AAF28547
 ID AAF28547 standard; DNA; 89047 BP.
 AC AAF28547;
 XX 04-APR-2001 (first entry)
 DE Genomic fragment #34.
 KW Genomic library; bacteria; human upper airway; otitis media; sinusitis;
 KW bronchopulmonary; endocarditis; meningitis; ss.
 OS Moraxella catarrhalis.
 PN W0200078968-A2.
 XX 28-DEC-2000.
 PD 16-JUN-2000; 2000WO-US16649.
 PP 18-JUN-1999; 99US-0140121.
 PR (INCY-) INCYTE GENOMICS INC.
 PA Lagace RE, Patterson C, Berg KL;
 PI WPI; 2001-041427/05.
 DR
 XX
 PT Genomic library for identifying diagnostic and therapeutic
 PT compositions, and for identifying virulence factors, regulatory
 PT elements and drug targets, comprises Moraxella catarrhalis nucleic
 PT acids -
 XX
 PS Claim 1; Page 324-345; 545pp; English.
 XX
 CC The present invention relates to a Moraxella catarrhalis genomic library
 CC comprising of a combination of 41 nucleic acid molecules (see
 CC AAF28514-AAF28554). The library has a number of uses described in the
 CC specification e.g. is useful for identifying diagnostic and therapeutic
 CC compositions. M. catarrhalis (Branhamella catarrhalis) is a large
 CC aerobic, gram-negative diplococcus, normally found among the bacterial
 CC flora of human upper airways. M. catarrhalis is known to cause acute,
 CC localised infections such as otitis media, sinusitis and bronchopulmonary
 CC infection and life-threatening, systemic diseases including endocarditis

CC and meningitis.

XX Sequence 89047 BP: 26501 A: 17338 C: 19060 G: 26147 T: 1 other:

Alignment Scores:

Pred. No.:	0	Length:	89047
Score:	4727.00	Matches:	919
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	22	Gaps:	0

US-09-914-168-2 (1-919) x AAF28547 (1-89047)

```
OY 1 MetSerLysProValLeuPheAlaAsnArgSerPheMetProValAlaLeuAlaIleTyr 20
DB 80397 ATGTCAAGCCGCTTTGTTTGGCAATCGAGTTTATGCTTCGCTTGGCGGCTTAT 80456
OY 21 LeuProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsn 40
DB 80457 TTGCTTTGATGACATCGCAACGATTGGCCACAACAAATTAACCTGCAACATCATCAAT 80516
OY 41 HisValProAlaHisAspThrAlaIleAsnGlnAlaLysAlaGlnLysnProProValLeu 60
DB 80517 CATGTACCCGCTGATACACCGCCATCATCAAGCAAGGAGGCAATCGCGCTTTTG 80576
OY 61 LeuThrProGlnGlnIleGlnAlaArgLeuAsnAlaIleGlnLeuAsnAlaLysProGln 80
DB 80577 CTAAACACCTGAGAGATACAGACCGCTTAATGCTGTGACTGAATGTAAGCCCAA 80636
OY 81 SerGlnAlaLeuAspValAlaAsnPheAspAspGlnSerProIleSerArgIleGly 100
DB 80637 TCACAGCTTTGGATGTCATTTTGTGATGATCAATCGCCGATATCTGTATCGGTGAG 80696
OY 101 GlnSerProProLeuGlnLysAspMetSerValIleGlnGlnThrProLeuSerLeu 120
DB 80697 CATACACCCCTTTGGCTTTGGATGTCGTGTCATGAGAAAGCAACACCGCTAACCTTG 80756
OY 121 GlnGlnLeuPheAlaGlnGlnSerThrGlnMetGlyIleAsnProAsnAspTyrIlePro 140
DB 80757 GAGGAAATTAATTTGGCTCAAACTACTGATGATCGGAATCAATCAATGATTATATATCA 80816
OY 141 GlnTyrGlnGlnGlnGlnProAsnSerGlnValAlaValProProThrLeuGlnProGln 160
DB 80817 GATATCAAGGCGGAGACCTAATAGTGTGTTGTACCCACGACATTAAGACCTGAA 80876
OY 161 LysProGlnLysLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspIleValAsnLysVal 180
DB 80877 AAACAGGTTTGTATCAAGCGCTTTATGCAACGCTTAATGATGTGTCAATAGTG 80936
OY 181 ProArgLeuLysAlaLysPheTyrGlnSerGlnSerGlnSerGlnSerAlaIleGly 200
DB 80937 CCTAGGCTTAAGGCAAAATTTTATCAATCATCGCAATCGGCAAAACCGTCCGATTGGG 80996
OY 201 SerSerHisGlnLysThrGlnProTyrAlaAsnIleLysAlaIleGlnLysPheThr 220
DB 80997 TCATGGCATCAAAAAACAGACGCTTATGCAATATCAAGCAACGCTTGAACATCAACC 81056
OY 221 GlnGlnSerAlaMetAspLeuAsnGlnSerIleProArgLeuArgGlnThrAlaLeuVal 240
DB 81057 CAAAGAGTCAGCATGATTTGAATGCTTATCCACGCTTAAGCAAACTGCTTTGGTG 81116
OY 241 AlaAlaArgAlaValAlaGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGly 260
DB 81117 GCAGCGCTGCTGCTGCTTATTATGATTTGATTTATCAATCATTAAGAAATAGCATCGGA 81176
OY 261 GluValAspValIleIleHisAspLeuGlnGlnProValTyrIleAspTyrArgAlaVal 280
DB 81177 GAGGTGATGATCATCATCATCATGATTAAGTGAACCTGTTTATTTGATTTATGACAGCGGTG 81236
OY 281 GluValAlaArgGlyGlnGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValPro 300
DB 81237 GAGGTACAGGTGAAGGTGCTGATGATTAAGCATTTACTACCGTGGCGATGAGGTGCCA 81296
```

```
OY 301 LeuLeuIleGlyAspValPheHisHisGlyLysTyrGlnThrLysLysAsnLeuIleGlu 320
DB 81297 TTGCTGATCGCGCATGTCTTTTCATCATGCGCAAGTACGAAACCAAAAAATTCATCATGAA 81356
OY 321 AsnAlaSerIleGlnHisGlyTyrPheAspGlyArgThrLeuAspArgSerValAspVal 340
DB 81357 AATGCCAGTCTCAACATGGATATTTTGTATGGGCGTGGCGATGCTTCAGTTGATGTA 81416
OY 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlnTyrArg 360
DB 81417 ATTTTGCAGATTAATACCGCTGATGCAGCTTAATTATGATACAGGTACCGCATTCGC 81476
OY 361 PheAspGluValAlaPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
DB 81477 TTGTAGAGGTGTATTTTATTTTACCATTTGATCTTAACCAATCAATTAATGCAACCGATCCA 81536
OY 381 AspLysLeuProValLysArgGlnLeuGlnGlnLeuLeuThrValAsnMetGlyGlu 400
DB 81537 GATAAGCTGCCAGTTAAAGAGAAATTAATTGACAGTACTACCGTTAAATGATGGAGAG 81596
OY 401 AlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
DB 81597 GCTTACAAATTTACAGCGGCTGCTGCATTTCAAAATGATTTGATTCACACAGGTATTTT 81656
OY 421 AsnMetValAsnThrGlnIleValPheProGlnArgGlnGlnIleGlnAsnAspGlnVal 440
DB 81657 AATATGTCGATACCGAGATTTGCTTTCCAGACGCTGAACAGATCCAAACGACCAAGTG 81716
OY 441 SerPheGlnGlnSerSerSerSerArgThrGlnProAlaGlnValAspLeuSerThrLeu 460
DB 81717 AGCTTTGACAGCTTCAAGTACGCTTACAGCAACGCAAGTTGATGAAAGCACACTT 81776
OY 461 GluProValIleGlnThrValGlnLeuThrAspGlyIleLeuMetAspIleSerProIle 480
DB 81777 GAACCTGTATTGAACCGCTTGAGCTACGAGATGGATTTAATGATATTTTCCGCCATC 81836
OY 481 GluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaLysAlaArg 500
DB 81837 GATTTAGTGCATCTTAATCTGATTCAAGACAGCTAAATTTGGTGGCTCCCAAGCTGCC 81896
OY 501 HisLeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
DB 81897 CATTTATATGACATGCTATATATAGTGTGCTTGCATCAATCATGATGATGCGCTAAT 81956
OY 521 ArgSerIleLeuGlnArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuPro 540
DB 81957 CGCTCATTTTGGGCAAGATCAAGCATGCCGTATCTGCCGTGCAAGTCTATTTTACTT 82016
OY 541 AspGlnSerGlnAsnGlnValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLys 560
DB 82017 GATGAATCTGAATAATGAGTATAGATTGGCCGAGCGTACGCAATTTGGCTAATGCAAG 82076
OY 561 ThrProAlaAspValTyrGlnSerLysValProLeuTyrValPheValAlaSerAsp 580
DB 82077 ACCCTGCTGATGTATCAAAAGTAAAAAGTCCGCTATATGTTCTTGTGGCGAGTGAT 82136
OY 581 LysProArgAspGlnGlnIleGlyLeuGlnTyrPglYSerAspThrGlyThrArgLeuVal 600
DB 82137 AAACACGAGATGTCAAATTTGTTGGCTTGGGATGCGACACAGTACCGGCTTAGTC 82196
OY 601 ThrLysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaLeuLeuArg 620
DB 82197 ACAAATTTGAGCATTAATTTGATTTATGCTGATGCTATCAAGCAGGCGCTAGAGCTAAGA 82256
OY 621 LeuSerGlnAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeu 640
DB 82257 CTGTTGAGGATAAAAAGGGCTCAAGTTATATGTCACCAAAACGCTTAAGCCACCTCTTA 82316
OY 641 AsnAspGlnLeuArgAlaThrLeuGlnTyrArgGlnGlnValPheGlyHisSerThrAsn 660
DB 82317 AATGATCAAGCTAAGAGCAACTTTGGTTATCAACAAGAAAGTTTTTGGTCACTCTTACCAAT 82376
```

Qy 661 GlyPheasPleuserThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsn 680
 Db 82377 GGTTCATATTATCCACACCGACCTAGAGCATGAGATTACCGCAGATTATATCCAAAT 82436
 Qy 681 GlyGlyTrpAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAla 700
 Db 82437 GGTGGCTGGAACTCTACTATTATTCATTGGCTTCATTGTAAGCTTAAACCCAAACAA 82496
 Qy 701 ProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGlu 720
 Db 82497 CCCCCCAACATGCGAGGATTTTACCACTGGATTTTGTCAATGGTAAGCCCAAGCCAAAG 82556
 Qy 721 AlaLeuAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyr 740
 Db 82557 GCGTTATTTGGCAGGTGGTGGCTGCATTAACAGGTTGCAGATTAATTGGTTAATCCGATG 82616
 Qy 741 ArgGlyIlyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAla 760
 Db 82617 CGTGGCATGCTCAGCAGATATCTTTAGAGGTTGGCTCAACGGTTGGTATCGGATGCT 82676
 Qy 761 AsnMetaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyr 780
 Db 82677 AATTATGCTATATGCTCCAGCTGGATATTAGTGGCTGTATAGTTTGGGATTAATGCTTAT 82736
 Qy 781 GlySerAsnArgAlaHisGlnMetThrGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
 Db 82737 GCGCAGCATGCTGCCCATCAGATCAGTGGCATACAGCAGATACATTGGTCGAT 82796
 Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820
 Db 82797 AATTATATCATGTCATATCCATATCGTTTGGCTTGGTGGCAGCAAGATTTGCT 82856
 Qy 821 GlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGln 840
 Db 82857 GGAATATGCACATGCAGTTTATCAGCTATATCAGATTAAGGTTATCTACAGCGCGTAA 82916
 Qy 841 ValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPheMetLysAspLeuArgLeuAla 860
 Db 82917 GTATTGGCGGTGGTACAGCTGAATATATGAAATTAAGAAAGATTTCCGTTTGGCG 82976
 Qy 861 ValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGly 880
 Db 82977 GTTTTGGTATATGTAATGCTTATGATGAAGGCTTACTAATGATGACCAAAATTTGCT 83036
 Qy 881 AlaGlyValGlyValArgTyrPalaSerProValGlyGlnValArgValAspValAlaThr 900
 Db 83037 GCAGCTGTCGCTGTCCTGGCGCATCAGCTGCTCAAGTTGCTTGAATGTGGCACT 83096
 Qy 901 GlyValLysGluGluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
 Db 83097 GGTGTCAAGAGAGGCAATCCATTAAGCTGATTTTATTTGGCACACCATTTT 83153

RESULT 3
 AAAS0537
 ID AAAS0537 standard; DNA: 2670 BP.
 AC AAAS0537;
 XX
 DT 05-DEC-2000 (first entry)
 DE Moraxella catarhalis mature protein coding region.
 XX
 KW BASB081 gene; infection; otitis media; pneumonia; sinusitis;
 KW inflammation; therapy; antibacterial; antiinflammatory; vaccine;
 KW diagnosis; ds.
 XX
 OS Moraxella catarhalis.
 XX
 FH Key location/Qualifiers
 FT conflict replace(1003,G)
 FT /tag- a
 XX /note- "corresponds to base 1093 of AAAS0536"

PN W0200052042-A1.
 XX
 PD 08-SEP-2000.
 XX
 PF 23-FEB-2000; 2000WO-EP01468.
 XX
 PR 26-FEB-1999; 99GB-0004559.
 XX
 PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
 XX
 P1 Ruelle J;
 XX
 DR WPI: 2000-587296/55.
 DR P-PSDB: AAY955988.
 XX
 PT New BASB081 polypeptides from Moraxella catarhalis and polynucleotides
 PT encoding the polypeptides used for treating infections, or as a vaccine
 PT for preventing infections, especially those caused by M. catarhalis -
 XX
 PS Claim 8; Page 60-61; 97pp; English.
 XX
 CC The present sequence is that of DNA encoding BASB081 mature protein
 CC (see AAY955988) of Moraxella catarhalis Mc2931 (ATCC 43617). It was
 CC obtained by sequencing PCR-amplified DNA from a BASB081 expression
 CC plasmid. The sequence shows 99.9% identity to a mature BASB081
 CC coding sequence (see AAAS0537) obtained from M. catarhalis ATCC 43617.
 CC The invention provides BASB081 polypeptides, polynucleotides,
 CC expression vectors, host cells, and a process for producing a BASB081
 CC polypeptide. Also provided are vaccine compositions comprising a
 CC BASB081 polypeptide or polynucleotide, and optionally at least 1
 CC other M. catarhalis antigen. A method for diagnosing a M.
 CC catarhalis infection involves identifying a BASB081 polypeptide, or
 CC an antibody that is immunospecific for it, in a sample. A
 CC therapeutic composition useful in treating M. catarhalis diseases
 CC in humans comprises an antibody directed against a BASB081 protein.
 CC The disease can be a bacterial infection, e.g. otitis media in
 CC infants and children, pneumonia in elderly, sinusitis, nosocomial
 CC infections and invasive diseases, chronic otitis media with hearing
 CC loss, fluid accumulation in the middle ear, auditive nerve damage,
 CC delayed speech learning, upper respiratory tract infection, and
 CC inflammation of the middle ear.
 XX
 S0 Sequence 2670 BP; 757 A; 527 C; 632 G; 754 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 0 Length: 2670
 Score: 4581.00 Matches: 888
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.89% Mismatches: 0
 Query Match: 96.91% Indels: 0
 DB: 21 Gaps: 0

US-09-914-168-2 (1-919) x AAAS0537 (1-2670)
 Qy 31 GlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisAspThrAlaIleAsn 50
 Db 1 CAACAAATAACCTCGCAACATCATCATGATACCCGCTCAGACACCGCCATCATAT 60
 Qy 51 GlnAlaLysAlaGlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaArgLeu 70
 Db 61 CAAGCAAGAGCAGGCAATCCGCTGTTTGCACACCTGAGACAGATACAAACACGCTT 120
 Qy 71 AsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnAsp 90
 Db 121 AATGCTGCTGGACGTGAATGCTTAAGCCCAATCACAAGCTTTGGATGTTGTCATTTTAT 180
 Qy 91 AspGlnSerProIleSerArgIleGlyLysGluGlnSerProProLeuGlyLeuAspMetSer 110
 Db 181 GATCAATCGCGGATATCTCGATATCGTGAGCAATCAACCCCTTGGTTGGATATGTCG 240
 Qy 111 ValIleGluGluThrThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGlu 130
 Db 241 GTCATCGAAGAAACACACCGCTTAAGCTTGGAGGAATTAATTGTCAGAAATCTACTGAG 300

QY	131	MecUyIleasnProasnAspTyrIleProGluTyrGlnGlyGlnInProasnSergIu	150
Db	301	ATCGCAATCAATCCAAATATGATATTATTCGCAAGATATCAAGCGCAACCTCAATAGTGAG	360
OY	151	ValValValProProPthrIleGluProGluLysProGluLysIleLysArgIleuTyrAla	170
Db	361	GTGGTTGTACACCGCACATTTAGACCTGCAAAAACAGGTTTGATCAAGGCTCTTATGCA	420
OY	171	ArgIleuPheasnAspGlyValAsnLysValProArgIleuLysAlaLysPheTyrGlnSer	190
Db	421	CGCCATTATTAATGATGCTGTCCAAATAAAGCTGCTTACGCTTAAGCAAAATTTATCATCA	480
OY	191	SerGlnSerGlyIuThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAla	210
Db	481	TGCGAATCAGCGCAACACGACGATGGCTGCTCATCAACAAAAACAGACGCTTATGCA	540
OY	211	AsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAlaMetAspLysnGlySer	230
Db	541	AATATCAACAGCAGCATTTGAAGACATCACCCAGACATCCAGCATGATTTGAATGGCTCT	600
OY	231	IleProArgIleuArgIlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIle	250
Db	601	ATCCACCGCCCAAGCGAAACCTGCTTGGTGGCGCCGCTGCTGCTGCTTATATCATATT	660
OY	251	AspLysSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspGluGly	270
Db	661	GATTATTCATCATCAACAATTACCATCCGAGAGCTGGATGTATCATCTCATCATTTAGCT	720
OY	271	GluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGluValAspAspLys	290
Db	721	GAACCTGTTTATATGATTTATTCAGCGCGTGGAGCTACGAGTGAAAGGTCTCATATATAA	780
OY	291	AlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisIleGly	310
Db	781	GCATTTACTACCGTGGCGGATGAGGTGCCATGTGCTATGGGAGATGCTTTTATCATCGGC	840
OY	311	LysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGlnHisGlyTyrPheAsp	330
Db	841	AAGTACGAAACCAAAAAAATCTCATCGAANAATGCCAGTGTGAACATGATATTTGAT	900
OY	331	GlyArgArgPheLysAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSer	350
Db	901	GGGCGTTCGCGGATGCTCTAGTTGATGTAATTTTGGCCAGTAATAACCGCTCATCTAGC	960
OY	351	LeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIleAsp	370
Db	961	TTAATTTATGTCACAGGTACCGCATGCTGTTGATGACAGCTATATTTTACCTATGAT	1020
OY	371	ProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeu	390
Db	1021	CCTTAAACCAATCAATTGACAAACCGATCAGATTAACCTGCCAGTTAAACGAAATTTACTT	1080
OY	391	GlnGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeu	410
Db	1081	GAGCAGTTACTCAACCGTTAAATCAGGAGAGCGCTTACAATTTACAGCGGCTGCTGACTT	1140
OY	411	SerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhePro	430
Db	1141	TCAATATGATTTGATTTGGCACACGATTTTAAATATGCTGAATACGACATGCTCTTTCCA	1200
OY	431	GluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArgThr	450
Db	1201	GAGCGTGAACAGATCCAAAACGCAAGTAGAGCTTTGACCACTTTCAAGTACCGGTACT	1260
OY	451	GluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThr	470
Db	1261	GAACACACACAATGTGATGAACACACATTTGAACCTGTCATTGAACGCTTGAGCTAACG	1320
OY	471	AspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAsp	490
Db	1321	GATGGGATTAATTAATGATATTTCGCCCATCGAATTTAGATCATCTAAATCTGATTCAGAC	1380

QY	491	LysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTrpAspMetProAspAspArgVal	510
Db	1381	AAAGCTAAATTTGGTGGCCCAAGGCGCCGACATTTATATACATGGCTGTGATAGGGTG	1440
QY	511	LeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAla	530
Db	1441	CTTGGCATCATCATGATGATGAGGGGTAAATCGCTTATTTTGGGAGGATCACCGATGCC	1500
QY	531	ValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeu	550
Db	1501	GTATGTGCGGTGGACAGCGCTATTTTACCTGATGAATCTGAAATAGAGTAATAGATTGG	1560
QY	551	ProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValArgGlnSerLysLys	570
Db	1561	CCCGAGCGCTACCGCATTTGGCTAATCGCAAGACCCCTGCTGATGTTCTATGAAGTAAAAA	1620
QY	571	ValProLeuTrpValPheValAlaSerAspLysProArgAspGlyGlnIleLeuGly	590
Db	1621	GTGCGCGCTATATGTCCTTTGTGGCAGATGAATAAACCCGAGATGTCAAATTTGGTTGGCG	1680
QY	591	TrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArg	610
Db	1681	TGGGATGTGGACACAGGATACCCGCTTAGTCACAAATTTGAGCATATTGATTATATCGT	1740
QY	611	AspGlyTrpGlnAlaGlyAlaGluLeuAsnArgLeuSerGluAspLysLysGlyValLysLeu	630
Db	1741	GATGCTATCAACAGCAGGGCGGAGACTAAGACTGTGAGAGATGAAAAAAGGGGTCAAGTTA	1800
QY	631	TyrAlaThrLysProLeuSerHisProLeuAsnAspGluAsnArgAlaThrLeuGlyTrp	650
Db	1801	TATGCCACCAAAACCGCTTAGCCACCCCTCTAAATGATCAGTTAAGACCAACTTTGGGTTAT	1860
QY	651	GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu	670
Db	1861	CAACAGAAAGTTTTGGTGCACCTGACCAATAGTTTTGATTATCCACAGCACCTAGAG	1920
QY	671	HisGluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThrTrpSerLeuArg	690
Db	1921	CATGAGATTACCCGACAGATTATATCCAAAAGTGCGTCGATCGTACTTATTCATTGGCGT	1980
QY	691	TyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGluAspLeuProVal	710
Db	1981	TATGCTCTGTATACCTTAAACCCCAAGCACCCCTGGAACATGCGAGATTTACCACTG	2040
QY	711	AspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys	730
Db	2041	GATTTTGTCAATGGTAAAGCCCAAGCAAGAGCGCTTATTGGCAGGGTCTCTGTCAATAA	2100
QY	731	ThrValAlaAspAsnLeuValAsnProMetArgGlyThrArgGluArgTrpSerLeuGlu	750
Db	2101	ACGGTTCAGATATTGTTGGTTAATCCATGCGGGCTATGCTGACGATATCTTTAGAG	2160
QY	751	ValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSer	770
Db	2161	GTTGGCTCAACCGGTTGGTATGATCGATGCTAAATATGGCTATTGCTCGACCTGTTATAGT	2220
QY	771	GlyValTrpSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGly	790
Db	2221	GGCGGTATAGTTTGGGGATAATAGCTTATAGCAGCAATCGGCCCATGACATGACTGGT	2280
QY	791	GlyIleGlnAlaGlyTrpIleTrpSerAspAsnPheAsnHisValProTrpArgLeuArg	810
Db	2281	GGCATACACAGAGATCATTTGGTGGGATTAATTTAATCATCATGTCGCAATTCGTTGGCT	2340
QY	811	PhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIle	830
Db	2341	TTTTTTCTGTGGCGACCAAGATTTCGTGGATATGACACATGACAGTTTATCACACTTA	2400
QY	831	SerAspLysGlyTrpLeuThrGlyGlnGlnValLeuAlaValGlyThrAlaGluTrpAsn	850
Db	2401	TCAATATAGGGCTTATCTGCACAGCGGCTCAAGTATTTGGCGGTGGTACAGCTGAATATAT	2460
QY	851	TyrGlnPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTrpAsp	870

|||||
 Db 2461 TATGAATTTATGAAAGATTTCGCTTGGCGCTTTTGGTGAATATGTAATGCTTATGAT 2520
 Qy 871 LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValArgTyrAlaSerPro 890
 Db 2521 AAAGGCTTACTAGATGATACCAAAATGTCAGGTGCGGTGCTGGCAGCACCT 2580
 Qy 891 ValGlyGlnValArgValAspValAlaThrGlyValLysGluGluGlyAsnProIleLys 910
 Db 2581 GTGGGTCAAGTTCGTGTGATGTGGCACTGTGTCAAGAAGAGGCAATCCCATTAAG 2640
 Qy 911 LeuHisPhePheIleGlyThrProPhe 919
 Db 2641 CTCATTTTATTTATGGCACACCATTT 2667
 RESULT 4
 AAF61044
 ID AAF61044 standard; DNA; 1374 BP.
 AC AAF61044:
 XX 16-MAY-2001 (first entry)
 DT
 XX
 DE P. putida KT2440-associated DNA ORF06604.
 KM
 KW Transgenic plant; detection; probe; amplification; vaccine carrier;
 XX microbial production strain; biological remediation; ds.
 OS Pseudomonas putida.
 XX
 PN DE1993088-A1.
 PD
 XX 01-FEB-2001.
 PF 27-JUL-1999; 99DE-1035088.
 PR 27-JUL-1999; 99DE-1035088.
 XX
 PA (TIGR-) TIGR INST GENOMIC RES.
 PA (OUIA-) OUIAGEN GMBH.
 PA (GBFB-) GES BIOECHANOLOGISCHE FORSCHUNG MBH.
 PA (DKFZ-) DKFZ DEUT. KREBSFORSCHUNGSZENTRUM.
 PA (MED1-) MEDIZINISCHE HOCHSCHULE HANNOVER.
 XX
 DR WPI; 2001-192469/20.
 XX
 PT New DNA sequences specific for Pseudomonas putida KT2440, useful as
 PT psf genetic engineering host, allow detection in presence of other
 PT related bacteria .
 XX
 PS Claim 1a; Page 94; 158pp; German.
 XX
 CC This invention describes novel DNA sequences (1) for specific detection
 CC of Pseudomonas putida KT2440. The invention also describes (1)
 CC recombinant expression vector containing (1); (2) prokaryotic or
 CC eukaryotic cells transformed or transfected with (1) or the vector of
 CC (1); (3) production of expression products by culturing cells of (2);
 CC (4) expression products, or their fragments, of (1) and synthetic
 CC proteins or peptides with the same sequences (A); (5) poly- or
 CC monoclonal antibodies (Ab) that react specifically with (A); (6)
 CC hybridoma cells that produce the monoclonal Ab of (5); (7) transgenic
 CC plants that contain transformed or transfected cells of (2); (8)
 CC detecting KT2440 using a labeled (1) or Ab as probe; and (9) DNA chips
 CC carrying one or more (1). (1), and their fragments, are used as probes
 CC to detect and isolate full-length cDNAs and/or to amplify such cDNAs by
 CC polymerase chain reaction, and for production of transgenic plants. (1),
 CC or antibodies that recognize their expression products, are used for
 CC detecting the presence of KT2440, particularly in presence of other,
 CC even closely related, bacteria. KT2440 is one of the bacteria classified
 CC work, e.g. as microbial production strains, for genetic engineering
 CC and as vaccine carriers. (1) are exclusive to KT2440 with no significant
 CC homology with sequences in other bacteria (specifically the closely

CC related pathogen P. aeruginosa). Compared with other 'safe' bacteria, it
 CC has greater catabolic activity and better survival in, and adaptation to,
 CC the rhizosphere and soil.
 XX
 SQ Sequence 1374 BP; 264 A; 428 C; 429 G; 251 T; 2 other; *
 Alignment Scores:
 Pred. No.: 6,84e-40 Length: 1374
 Score: 604.50 Matches: 165
 Percent Similarity: 40.41% Conservative: 90
 Best Local Similarity: 26.15% Mismatches: 195
 Query Match: 12.79% Indels: 181
 DB: Gaps: 12
 US-09-914-168-2 (1-919) x AAF61044 (1-1374)
 Qy 290 LysAlaPheThrThrValAlaAspGluValProLeuIleGlyAspValPheHisHis 309
 Db 4 AAGGCTTC---CGCGTGCCTGACAGCAAGCGCGCTGCGCGGTGACACATCAATCAC 60
 Qy 310 GlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPhe 329
 Db 61 GGCCATTACKAGGATGTCAGCGGTTGATCCAGAACCGAGCGTCGCGCTATGCTTCTTC 120
 Qy 330 AspGlyArgTyrPheAspArgSerValAspValIleLeuProAspAsnThrAlaAspVal 349
 Db 121 AGTGGCGGCTTACAGTACAGCGCGCTGCGCCGTGACGCCGACAGCGGTGCGGTATG 180
 Qy 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhePheThrIle 369
 Db 181 GAACGTGCTTACAGAGAGGCGCGGCTTATCCGTCGCGCGCTCACCTTC----- 231
 Qy 370 AspProLysThrAsnGluLeuThrThrAspProAspLysLeuProValLysArgGluLeu 389
 Db 232 -----GGTGGCGACACACCGCTGACGAGACCTG 261
 Qy 390 LeuGluGlnLeuLeuThrValAlaSerMetGlyGluAlaTyrAsnLeuGlnAlaValArgAla 409
 Db 262 CTCGACGCGATGTCGTGCTTCAAGCGGCTTACCCCTTACAGCTCGGAACTGTCGACAG 321
 Qy 410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhe 429
 Db 322 CTCGACCAACGACCTCGAATCGAGCGGCTATTTC----- 354
 Qy 430 ProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArg 449
 Db 354 ----- 354
 Qy 450 ThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeu 469
 Db 354 ----- 354
 Qy 470 ThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGln 489
 Db 355 ---GAAGGCGTCGCGGTGATGCGGCGCC----- 381
 Qy 490 AspLysLeuAsnLeuValAlaIleLysAlaIleArgHisLeuTyrAspMetProAspAspArg 509
 Db 381 ----- 381
 Qy 510 ValLeuAlaIleAsnHisAspAspGlyValAlaAsnArgSerIleLeuGluYrGluLeuSerAsp 529
 Db 381 ----- 381
 Qy 530 AlaValSerAlaValAlaIleArgAlaIleLeuProAspGluSerGluAsnGluValIleAsp 549
 Db 381 ----- 381
 Qy 550 LeuProGluArgThrAlaLeuAlaLeuAlaSerArgLysThrProAlaAspValTyrGlnSerLys 569
 Db 382 -----ACTGTGCTCTGCGCGAGAGAGTCCCGGCGAGTGTAT----- 420
 Qy 570 LysValProLeuTyrValAlaPheValAlaSerAspLysProArgAspGluGlnIleGluLeu 589

```

Db 421 -----CTGAAACCGGTAACCCAGCAGCACCAGCGGCGCTGGCGCTG 459
OY 590 G1YTPG1YSerAPThrG1YThrArgLeuVal1ThrIysPheGlnHisAsnLeuIleAsn 609
Db 460 GCGTTCTGACGAGCGTGGCGCGCGGCAAGCCAACTGGACCCCGGCTGAGTCAAC 519
OY 610 ArgAspG1YTyrglnAlaG1YAlaG1YLeuArgLeuSerGlnAspIysGlyValIys 629
Db 520 CCACAAAGCCACAGCTATGCTGGCAACCAAGCTGCGCGCGCGCGCGCGCGCGCG 579
OY 630 LeuYrAlaThrIysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGly 649
Db 580 CTGGGTATGACATTCTCCCTGACCCCGCGCTGACCCGCAAGTGGCTTGGCGGCGG 639
OY 650 Tyrgln--GlnGluVal1PheG1YHisSerThraSngIYpheaSpleuSerThraArgThr 668
Db 640 TACCGAAGCAGAGAGCTTGGCGGCGACACAGCGTTCACAGCTTTCAGCGTGGCGCC 699
OY 669 LeuGlnHisGluIleSerArgSerIleIleGlnAsnGlyIYTPAsnArgThrTyriser 688
Db 700 GAGTGGCACACCAAGCTGCCCTG-----GGCTGGCACGCGGCTGATTTCG 744
OY 689 LeuArgTyrgLeuAspIysLeuIysThrcIAlaProProGlnIuThrTrpGlnAspLeu 708
Db 745 CTCAAGTACCAAGCGGAGAAATATGCTG----- 774
OY 709 ProValAspPheValAsnG1YIysProSer-----GlnGluAlaLeuLeuAlaGlyVal 726
Db 775 -----GGTGACGACTCCGCTTGGACCAACCTGCTGATCCGCGGCGCTC 816
OY 727 AlaValHisIysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrglnArg 746
Db 817 AGCTTTTCTCTTCCTGCGCAGGACACGATCGATCGATCGGCAACAGGCTATCGCTGAC 876
OY 747 TyrSerLeuGluVal1G1YSerSerG1YLeuVal1SerAspAlaAsnMetAlaIleAlaArg 766
Db 877 TTCATGTTCAAGGTGCCAAGGAAGGCTGCTGTCGACACCAACCTCTGACAGCGCAC 936
OY 767 AlaGlyIleSerGlyVal1YrSerPheG1YAspAsnAlaTyrgIYSerAsnArgAlaHis 786
Db 937 GTATTGCTCAAGGCGCTGACCACTCGGCCAACAC-----CAC 975
OY 787 GlnMetThrGlyGlyIleGlnAlaGlyIYrIleTrpSerAspAsnPhe--AsnHisVal 805
Db 976 CGCTTTCTCGGACGCGATCTTGGTGGCAAGTGCACCAACCAATGGCTTCAGAACACACTT 1035
OY 806 ProTyrgLeuArgPhePheAlaG1YIysAspIleSerIleArgIYTyrgAlaHisAsp 825
Db 1036 CCGCGCTGCGCTGCTTCTTGGCGGTGGCGACCAAGATGCGCGGTACGACTACGAC 1095
OY 826 SerLeuSerProIleSerAspIysG1YrLeuThrcIYsGlnVal1LeuAlaValGly 845
Db 1096 ACCCTGTCGCGAAGACAGCAGCAGCTGATCGCGCGGCGCTACTTGTGTCGACAGC 1155
OY 846 ThrAlaGluTyrgAsnTyrgIYrPheMetIysAspLeuArgLeuAlaVal1PheG1YAspIle 865
Db 1156 AGTGTGCGAGTACAGATATCTCTGACGAAATAATGGCGGCGGAGCTGTGCTGCAAA 1215
OY 866 GlyAsnAlaTyrgAspIysG1YrPheThraSnpThrIysIleGlyAlaG1YValGlyVal 885
Db 1216 GGCAGATCGCTTCAAGAGCTGAGCTGCGCGCGCGCGCGCGCGCGCTGGTGGTGG 1275
OY 886 ArgTrpAlaSerProVal1G1YGlnVal1ArgVal1AspVal1AlaThrcIYsVal1GlyGlu 905
Db 1276 CGGTGGGTATGCGCAGTGGCGGCGCTGCGCTGAGCAGCTGCGCAAGCGCTGATGACGAA 1335
OY 906 GlyAsnProIleIysLeuHisPhePheIleGly 916
Db 1336 GGGGCG---ATTGCGCTGCACTTTTTCATGGGG 1365

```

```

ID AAA50269 standard; DNA: 1737 BP.
XX
AC AAA50269;
XX
DT 07-NOV-2000 (first entry)
XX
DE Haemophilus influenza strain Rd KW20 BASB067 gene.
XX
KW BASB067 gene; outer membrane protein; vaccine; antidiotic;
XX antibacterial; screening; infection; diagnosis; therapy; ds.
XX
OS Haemophilus influenzae.
XX
FH Key Location/Qualifiers
FT CDS 1..1737
FT FT /*tag= a
FT sig_peptide 1..66
FT FT /*tag= b
FT mat_peptide 67..1734
FT FT /*tag= c
XX
PN MO200047737-A1.
XX
PD 17-AUG-2000.
XX
PE 04-FEB-2000; 2000WC-EP00887.
XX
PR 09-FEB-1999; 99GB-0002880.
XX
PA (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Ruelle J, Thonnard J;
XX
DR WPI: 2000-515059/46.
DR P-PSDB: AAY95820.
XX
PT BASB067 polypeptide and polynucleotide from Haemophilus influenzae are
PT used for diagnosing and treating H. influenzae infections -
XX
PS Claim 3; Page 77-78; 87pp; English.
XX
CC The present sequence is that of the coding region of the BASB067
CC gene of Haemophilus influenzae strain Rd KW20. It encodes a
CC 578-amino acid surface expressed protein (see AAY95820) that is
CC recognized by the immune system. The invention relates to
CC recombinant materials and methods for the production of BASB067
CC polypeptides and polynucleotides, for use especially in therapeutic
CC and prophylactic vaccines. It also relates to methods for using
CC such polypeptides and polynucleotides in the prevention and
CC treatment of microbial diseases, in diagnostic assays for detecting
CC diseases associated with microbial infections, and assays for
CC detecting expression or activity of BASB067 polypeptides or
CC polynucleotides. A polynucleotide having at least 85% identity
CC to the present sequence can be used in the recombinant production
CC of BASB067 immunogenic polypeptides in transformed host cells, and
CC in vaccine compositions.
XX
SO Sequence 1737 BP; 517 A; 309 C; 380 G; 531 T; 0 other;
XX
Alignment Scores:
Pred. No.: 4.6e-26 Length: 1737
Score: 436.50 Matches: 142
Percent Similarity: 35.64% Conservative: 105
Best Local Similarity: 20.49% Mismatches: 235
Query Match: 9.23% Indels: 211
DB: Gaps: 14
US-09-914-168-2 (1-919) x AAA50269 (1-1737)
OY 243 ArgAlaValGlyTyrgIYrAspIleAspLeuSerIleIleArgAsnSer---IleGlyGlu 261
Db 217 CGTGTGTTGGTATATGAATCTTCGCTGGCTTTCAGACGAAACAGCGTCAAGGCGAA 276

```

QY 262 ValAspValIleIleHisAspLeu-----GlyGluProValIleAspTyrArgAla 279
 |||:|||||
 Db 277 CGGATTATTGATTGATTCATGATACACGCGGACCAACAAAATTTGGGGGACGTAT 336
 QY 280 ValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluVal 299
 |||:|||||
 Db 337 GTCCAAATTTAGGGGAGCGCGCACAGATGAAAATTTTAAATGATCCGTAACAACTTA 396
 QY 300 ProLeuLeuIleGlyAspValPheHisIleGlyLysTyrGluThrLysLysAsnLeuIle 319
 |||:|||||
 Db 397 CCG---AAAGATGCCGTTTGTTGGTGGACCAACCTTACGATGATTACAAACACGCGATT 453
 QY 320 GluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPheLysAspArgSerValAsp 339
 |||:|||||
 Db 454 TCACGTTTACGATTAAATCGTGGGATTTTGTATGGGAACTTAAATTTACGTTTAA 513
 QY 340 ValIleLeuProAspAsnThrAlaAspValSerIleTyrAspThrGlyThrGlyTyr 359
 |||:|||||
 Db 514 ATCAGCCCTGAAACCCATCAAGCATGTGGCGAATGTTATTTGATGATGTCCTCAT 573
 QY 360 ArgPheAspGluValValPhePheThrIleAspProLysThrAsnGluLeuThrThrAsp 379
 |||:|||||
 Db 574 CATTATGGCAATATTACTTT----- 594
 QY 380 ProAspLysLeuProValIleArgGluLeuLeuGluLeuLeuThrValAsnMetGly 399
 |||:|||||
 Db 595 ---AGCCATTACAAATTTGGGATGATTAATAATATATTCTTAACATCAAAATCTGCT 651
 QY 400 GluAlaLysAsnLeuGluAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyr 419
 |||:|||||
 Db 652 GATCCGATTTAAATGAATTAATTTGCGATTTAACACAGCGATTTCCATCTTCAAAATGG 711
 QY 420 PheAsnMetValAsnThrGluIleValPheProGluArgGluGluIleGluAsnAspGln 439
 |||:|||||
 Db 712 TTATGCTCAGTA----- 723
 QY 440 ValSerPheGluGlnSerSerSerArgThrGluProAlaGluValAspGluSerThr 459
 |||:|||||
 Db 723 ----- 723
 QY 460 LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro 479
 |||:|||||
 Db 723 ----- 723
 QY 480 IleGluPheSerAlaSerAsnLeuIleGluAspLysLeuAsnLeuValAlaAlaLysAla 499
 |||:|||||
 Db 724 -----TAGTTCAGCCCTAATGTTAAT----- 744
 QY 500 ArgHisLeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyVal 519
 |||:|||||
 Db 744 ----- 744
 QY 520 AsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeu 539
 |||:|||||
 Db 744 ----- 744
 QY 540 ProAspLysSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArg 559
 |||:|||||
 Db 744 ----- 744
 QY 560 LysThrProAlaAspValIleTyrGlnSerLysLysValProLeuTyrValPheValAlaSer 579
 |||:|||||
 Db 745 -----CATAAAGCAAACTGTGATGTGCGAGATTAATCTTTATCA 786
 QY 580 AspLysProArgAspGlyGluIleGlyLeuGlyTyrPheLysSerAspThrGlyThrArgLeu 599
 |||:|||||
 Db 787 CGTAAATAAAATGGGATGGAACCTGCTGGCTTTCTACTGATGGCGGCTTCACCGA 846
 QY 600 ValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyValGluLeu 619
 |||:|||||
 Db 847 CAATATAGCTGACAAACCTTGATTAATAGCCGTGACACATAGTTGCCGTTCAAAATCTT 906
 QY 620 ArgLeuSerGluAspLysLysGlyValLys--LeuTyrAlaThrLysProLeuSerHis 638

Db 907 TATCTCTGACCAACAAACCTAGAGGCACTATCAATGCACCTGTTAAATAT 966
 |||:|||||
 QY 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658
 |||:|||||
 Db 967 CCATTAAATTTTACTATGATTTTCCGCTGGTTGGGAA-----GGGGAATAA 1014
 QY 659 ThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIle 678
 |||:|||||
 Db 1015 GAGAAAT-----GATACCAATTCGAGAGTCTACGTTGTCAGCGTTACGTTATTTGGAAT 1068
 QY 679 GluAsnGlyGlyTyrPAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThr 698
 |||:|||||
 Db 1069 AATGGCATGGTGGCAATATTTTGGCGACTTCGTATGCGATACGACAGTTT---ACA 1125
 QY 699 GlnAlaProProGluThrThrGlnAspLeuProValAspPheValAsnGlyLysProSer 718
 |||:|||||
 Db 1126 CAAAGCGAT----- 1134
 QY 719 GlnGluAlaLeuLeuAlaGlyValAlaHisLysThrValAlaAspAsnLeuValAsn 738
 |||:|||||
 Db 1135 -----ATCAGTCAATAAACCTTA-----CTTCTTTAT 1161
 QY 739 ProMetArgGlyTyrArg----- 744
 |||:|||||
 Db 1162 CCAACTGTTGATTTACTGCTACTCGATTAACGTGGTGTTCTTGGCCACTTGGGGCGAT 1221
 QY 745 ---GlnArgTyrSerLeuGluValGlyLysSerGlyLeuValSerAspAlaAsnMetAla 763
 |||:|||||
 Db 1222 GTGCAAAATAATTAATCTTTGATTAAGCAAAACGAATTTGGCTVTCAGAAATCTTTTATA 1281
 QY 764 IleAlaArgAlaGlyIleSerGlyValIleSerPheGlyAspAsnAlaIleGlySerAsn 783
 |||:|||||
 Db 1282 AAAGTCAAGCATCTAAGCGCGGTGGCTGCTACTTATGCGAATAA----- 1326
 QY 784 ArgAlaHisGlnMetThrGlyGlyIleGluAlaGlyTyrIleTyrSerAspAsnPheAsn 803
 |||:|||||
 Db 1327 -----CATCGTGTGCTGCTGCTGCTGAAATCGGGTATTTTCAATACAAAGCTATTGCA 1380
 QY 804 HisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArgIleTyrAla 823
 |||:|||||
 Db 1381 AAAATTCGCGCTACACACTCGTCTTCTGCTGGTGGCGATGAGTGGCGGTTACGGC 1440
 QY 824 HisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAla 843
 |||:|||||
 Db 1441 TATATAAAATTTGCCCTTAAATAATAGAAATGCAAAATTTGCTGGCTTCCGATTTGCT 1500
 QY 844 ValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGly 863
 |||:|||||
 Db 1501 ACCACTTCTTTAGATATATCAATATATCAAGTCTATCCAAATTTGGGGCGCACACTTTTGA 1560
 QY 864 AspIleGlyAsnAlaIleTyrAsnTyrGlyPheThrAsnAspThrLysIleGlyAlaGlyVal 883
 |||:|||||
 Db 1561 GATAGTGAATTAAGCTGCGCATTAATTACACAGCAAAAGCACTGCTTATGCGCAGCGCTT 1620
 QY 884 GlyValArgTyrPAlaSerProValGlyGlnValArgValAspValAlaThrGlyValLys 903
 |||:|||||
 Db 1621 GGTGTGCGTGGCATTCGCGCAATTTGCGGATTAATTTGATATTTGCGCACACCATTCGT 1680
 QY 904 GluGluGlyAsnProIleLysLeuHisPhePheIleGly 916
 |||:|||||
 Db 1681 GATTAAGATTAACAGCAAAATATTCATTTTACATCCGA 1719
 RESULT 6
 ID AAA50270 standard; DNA; 1731 BP.
 AC AAA50270;
 AC XX
 AC XX
 DT 07-NOV-2000 (first entry)
 XX
 DE Haemophilus influenza non-typeable strain 289 BASB067 gene.


```

OY 659 ThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIle 678
DB 1009 GAGAAC-----GATACCATACAGAGAGTGCTTACGTTGCAGCGGTACGTTATGGCAAT 1052
OY 679 GluAsnGlyGlyThrPheAsnArgThrTyrSerLeuArgTyrArgLeuAspLysIleLysThr 698
DB 1063 AATGCCATGCGTTGGCAATATTTTGGCGGACTTCGTACGCGATACGACAGTTT---ACA 1119
OY 699 GluAlaProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysProSer 718
DB 1120 CAAGCGGAT----- 1128
OY 719 GlnGluAlaLeuLeuAlaGlyValAlaValHisTyrThrValAlaAspAsnLeuValAsn 738
DB 1129 -----ATCATCTGATTAACCTTA-----CTTCTTTAT 1155
OY 739 ProMetArgGlyTyrArg----- 744
DB 1156 CCAACTGTTCGATTACTCGCACTCGATTCAGTGTGCTTCCTTGGCACTTGGGCGAT 1215
OY 745 ---GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
DB 1216 GTGCAAAAATTTACTTTTGAATTAAGCAAAAGAAATTTGGCTATCAGATCTTCTTTATA 1275
OY 764 IleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsn 783
DB 1276 AAGTCGCAACATCTAGCGCGTGGTTCGACTTATCGCAAAAT----- 1320
OY 784 ArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsn 803
DB 1321 -----CATCGTATCGTTCGCTCGTGAATCGGATATTACATCAAAAAGATATTGAA 1374
OY 804 HisValProTyrArgLeuArgPhePheAlaGlyLysPheGlnSerIleArgGlyTyrAla 823
DB 1375 AAAATTCGCGCTACACTGCGCTTTCTTTGCTGTGGCATGCTAGTGTGGCGGTTACGGC 1434
OY 824 HisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAla 843
DB 1435 TATTAATAAATTTGCGCTAATAAATGAAATGAAATTTGCTGGCGCTCAGCTTGCCTT 1494
OY 844 ValGlyThrAlaGlyTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGly 863
DB 1495 ACCACTTCTTATGAAATCAATATCAAGTTTATCCAAATTTGTCGGCGCAACTTTGCA 1554
OY 864 AspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyVal 883
DB 1555 GATAGTGGATTACGTCGCGATTAATTACACAGCAAAAGAGCTGCTTATGGCACAGGCGTT 1614
OY 884 GlyValArgTyrPheAspProValGlyGlnValArgValAspValAlaThrGlyValLys 903
DB 1615 GGTGTGCGTTGGCGCATGCGCAGTGGCGATTAATTGATATATGCCACCCATTCGT 1674
OY 904 GluGluGlyAsnProIleLysLeuHisPhePheIleGly 916
DB 1675 GATTAAGATTAACAGCAAAAATATTCAATTTTACATCGGA 1713
RESULT 7
ID AAS93089 standard; cDNA; 2193 BP.
XX AAS93089;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #28893.
XX
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX
XX WO200175067-A2.

```

```

XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HXSE-) HXSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR P-PSDB; ABG28902.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 28893; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2193 BP; 481 A; 608 C; 582 G; 522 T; 0 other.
XX
Alignment Scores:
Pred. No.: 1.96e-16 Length: 2193
Score: 320.00 Matches: 176
Percent Similarity: 38.14% Conservative: 107
Best local Similarity: 23.72% Mismatches: 288
Query Match: 6.77% Indels: 171
DB: 23 Gaps: 30
US-09-914-168-2 (1-919) x AAS93089 (1-2193)
OY 237 ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIle 256
DB 2145 ACCGCGTAGTGGCGGCGCTGTAACGAGAAA-----GACCTCTGAGAGTTCAAC 2095
OY 257 AsnSerIleGlyGluValAspValIleIleHisAspLeuGlyLysProValTyrIleAsp 276
DB 2094 ATCGCAACTGGCCAGAGAGAGGTTCTCTTTTCACGCGCTTACGATCCGAATTCAGAAA 2035
OY 277 TyrArgAlaValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAla 296
DB 2034 GATCAGGCGCATGACCTCGCAACGAT----- 2008
OY 297 AspGluValProLeuLeuIleGlyAspValPheHisHisGlyTyrGluThrLysLys 316
DB 2007 -----CCGCTGCGAATGACGCGCTG-----AAAGACCGGCAACAAA 1969
OY 317 AsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAspArg 336

```

```

Db 1968 GCGAAATCGAACTGCTCCGCTCAG----- 1942
Oy 337 SerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGly 356
Db 1941 CAGACCGCGTTAACTCCCATATCATCTGACGACGCG-----ACCGGT 1897
Oy 337 ThrGlnTyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeu 376
Db 1896 CCGAAACACATGACATGAAATGACTGCTGGCAACCGCTGGTTGAAGATCTG 1837
Oy 377 ThrThrAspProAspLysLeuProValLysArgGluLeuLeuGlnIleLeuThrVal 396
Db 1836 GTAAACCGTTCATTGAG---CCGCTGAAAGTTGCATGCTGACGACGCTGGCTGCCGA 1780
Oy 397 AsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAla 416
Db 1779 TCTCAT-----ATCGACGACGTTATC--- 1759
Oy 417 ThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGln 436
Db 1758 -----CTCGTTGGTGGTCACTCGATCTGATCCCA-----ATGGTTGAG 1723
Oy 437 AsnAspGlnValSerPhe-----GluGlnSerSerSerSerArgThrGluProAla 453
Db 1722 AAGAAAGTTCTGAGCTCTTGGTAAAGAGCGCGCTAAAGACGTTAAACCGGACGAAGCT 1663
Oy 454 GlnValAspGluSerThrLeuGluProValIleGluThrValIleLeuThrAspGlyIle 473
Db 1662 GTAGCAATCGCGTGGTGGTTCAGGGTGGTCTTCGACTGCGTGAAGAGTAAACACGCTAG 1603
Oy 474 LeuMetAspIleSerProIleGluPheSerAlaSerAsnLeu----- 487
Db 1602 CTGCTGGACGTTACCCGCTGCTGCTGGTATGCAACCGCGGCTGTGATGACGACG 1543
Oy 488 IleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAsp 507
Db 1542 CTGATCGCGGAAAAACACACATATCCCGACCAAGCAGCAGAGTGTCTCTACCGCTGA 1483
Oy 508 AsparGlyValLeuAlaIleAsn---HisAspAspGlyValAsnArgSerIleLeuGlyArg 526
Db 1482 GACAACCGCTGCGGTAAACCATCCAT-----GTGCTGACGAGTGAACGTTAAACGT 1432
Oy 527 IleSerAspAlaValSerAlaVal-----AlaArgAlaIleLeuProAsp 541
Db 1431 CCGCGTGAATACAAATCTGCTGATACACTGGCAGCATACAGCCGGTGC----- 1384
Oy 542 GluSerGluAsn---GluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLys 560
Db 1383 ---ACGGAATATGTTCACTATATTGATACACCAGAGGGCTTATATATACGACCCCT 1327
Oy 561 ThrProAlaAspValTyrGlnSerLysLysValPro-----LeuTyrValPheValAla 578
Db 1326 CACCGTCGCCAACAGCCCACTAATCAACGCCCACCAACTCTCCACTGGGTGGCCGA 1267
Oy 579 SerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrProGlySerAspThrGlyThrArg 598
Db 1266 CGAGACCTAACCCGAACACCACCGCGCTGGACAGGACCCGAATAATTATGAGGCGC 1207
Oy 599 -----LeuValThrLysPheGluHis----- 605
Db 1206 CATGACACTATGCTCGCAAGGGGATTTGACGCGCTCCATACATGTTCCGGGCTTCCGA 1147
Oy 606 -----AsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSer 622
Db 1146 GAGAACCAGATCTCTGATCTGACTGCTACTCAGAAAGTTTCCGCTGGCTGAAGC 1087
Oy 623 ---GluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
Db 1086 GTTTCAGTCCGACCGGGCTA-----CCCGCTGCCGATTGTGAATCCGG 1042
Oy 642 AspGlnLeuArgAlaThr-----LeuGly 649

```

```

Db 1041 AGCAGTTGGTTGGACGCGCTCGTTATCATGACGCTGCTGATACCTTCATCTGCAT 982
Oy 650 TyrGlnGlnGluValPheGly-----HisSerThrAsnGlyPheAspLeuSer 665
Db 981 TTCAAAACGCGCGCGCTGGCAATCATGTCCTATTTGCCGAACGCGCATATGACTACT 922
Oy 666 ThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsn----- 680
Db 921 GCC-----TTCCGAACCCCGGCATACACAG 895
Oy 681 ---GlyGlyTyrAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGln 699
Db 894 CAAGGGCGGCGGAGCTGCATTAACCTGCGTGGAGTCTGCAC----- 850
Oy 700 AlaProGluThrThrProGlnAspLeuProValAspPheValAsnGlyLysProSerGln 719
Db 849 -----CACTTACTAGGCTGAATATTCACAT 823
Oy 720 GluAlaLeuLeu-----AlaGlyValAlaValHisLysThrValAlaAspAsnLeuVal 737
Db 822 ACCACGATGCTGTTATCTCGTGGGTGATGATTAGCCGACGCGTCTCGTGGGCGCTG 763
Oy 738 AsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValAlaLysSerGlyLeuVal 757
Db 762 ATGCCAAGCTGGGGCACTGCAACGCTACTCTATGCACTATGCAACAGCGCTGGCGT 703
Oy 758 SerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp 777
Db 702 TCAGATGCTGATTTTCCTTCCTCCAGGCCACAGCTGCGATCCGCACTGATGAT 643
Oy 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIle 797
Db 642 CGC-----CATCGTTTGTAAACGCGGCAAGCTGCTGGATTT 604
Oy 798 TrpSerAspAsnPheAsnHisValProTyrArgLeuAlaPhePheAlaGlyGlyAspGln 817
Db 603 GAACCGGTGATTTTCACAAGTACCGCGCGGATCTGGTGTCTTCGCGGGGCGACCGC 544
Oy 818 SerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThr 837
Db 543 AGTATTCGTGGGTACAAATATCTATGCTGCCGAATATCCCAACGAGGTGACTGAAA 484
Oy 838 GlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPheMetLysAspLeu 857
Db 483 GGGGCTGCAAGTTGATTAACCGGATGCTGGAATACCAAGTACCAAGTACCGGAAAAATG 424
Oy 858 ArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThr 877
Db 423 TGGGGCGCGGTGTTTGCATAGTGGCGAAGCGGTAAAGCATATTGCGCGGACGACTTT 364
Oy 878 LysIleGlyAlaGlyValAlaArgTyrAla-SerProValGlyGlnValArgValAs 897
Db 363 AAAACCGGTACCGCGGCGGCGCTGGGAATGGCGGCTGCGGCAATCAACTCGCA 304
Oy 897 PValAlaThrGlyVal-----LysGlnGluGlyAsnProIleLysLeuHisPheIle 915
Db 303 TTTTGGCGTACCGCTGCGGATTAAGACGACACGCG-----TTACAGTTTATCAT 253
Oy 915 eGly 916
Db 252 CGGT 249

```

RESULT 8
 ABR52155
 ID ABR52155 standard; DNA; 944 BP.
 XX
 AC ABR52155:
 XX
 DT 13-AUG-2002 (first entry)
 XX
 DE Haemophilus influenzae strain BASB207 upstream nucleotide sequence.
 XX
 KW BASB207; antibacterial; anti-inflammatory; vaccine; ss;


```

FT      /tag" a
FT      /product= "N. meningitidis serogroup A antigen"
FT      sig-peptide
FT      1..63
FT      mat-peptide
FT      64..2376
FT      /tag" b
FT      /tag" c
FT      /product= "Mature N. meningitidis serogroup A antigen"
PN      WO200138350-A2.
XX      31-MAY-2001.
XX      28-NOV-2000; 2000MO-IB01851.
XX      29-NOV-1999; 99GB-0028197.
XX      09-MAR-2000; 2000GB-0005696.
XX      (CHIR-) CHIRON SPA.
XX      (STAT-) STATENS INST FOLKEHELSE.
XX      Glutlan1 MM, Pizsa M, Rappuoli R, Holst J;
XX      MPI: 2001-381289/40.
XX      P-PSDB: AAU03958.
XX      Novel 85 kDa antigen from Neisseria meningitidis and Neisseria
XX      gonorrhoeae, useful in the manufacture of a medicament for treating and
XX      preventing Neisserial bacteria infection -
XX      PS      Claim 6: Flg 2: 92pp: English.
XX
XX      The sequence represents a DNA encoding a Neisseria meningitidis serogroup
XX      A 85 kDa antigenic protein. Neisseria meningitidis colonises the pharynx,
XX      causing meningitis and, occasionally, septicaemia in the absence of
XX      meningitis. This antigenic protein is useful in the manufacture of a
XX      medicament for treating or preventing infection due to Neisseria
XX      bacteria, such as meningitis and septicaemia. It is also useful as a
XX      diagnostic reagent for detecting the presence of Neisseria bacteria or
XX      antibodies raised against Neisseria, and as a reagent for raising the
XX      antibodies. The Neisserial nucleotide sequences can be expressed in a
XX      variety of different expression systems, for example, mammalian cells,
XX      baculoviruses, plants, bacteria and yeast.
XX
XX      Sequence 2379 BP; 636 A; 755 C; 596 G; 392 T; 0 other:
XX
Alignment Scores:
Pred. No.: 6.03e-12 Length: 2379
Score: 265.50 Matches: 195
Percent Similarity: 34.288 Conservative: 120
Best Local Similarity: 21.228 Mismatches: 348
Query Match: 5.628 Indels: 257
DB: 22 Gaps: 39
US-09-914-168-2 (1-919) x AAS07278 (1-2379)
QY      116 ThrProLeuSerLeuGluGluLeuPheAlaGluGluSerThrGluMetCylLeuAsnPro 135
DB      46 TCGCCTTTTGGCATTTGCCGACTTTCACCATCCGATCCGTCGCAAGCGTTGCGAGCGT 105
QY      136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGluInProAsnSer 149
DB      106 ACCGAGCCGAGCAGCGATTCACACTACGCGCCGTCACAAAGTCGGCAGCACCTACAAAC--- 162
QY      150 GluValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
DB      163 -----GACACACACGCGCAGTCGATCATCAAAAGCGCTGTAC 198
QY      170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
DB      199 GCCACCGCTTCTTTCGACGACGTCAGACGCAAACTCGCGAGCGGACGCTCTGTGACCC 258
QY      187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206

```

```

DB      259 GTTATGCAAGCCGCCACCATCGGCTCGCTCAACATCACCGGCGCCAAATGCTGCAAAAC 318
QY      207 GluProTyrAlaAsnIleLysAlaAlaLeuGluAsp-----IleThrGlnGluSerAla 224
DB      319 GACGCGC-----ATCAAGAAAAMAACTCGAAATCGTTGGGGCTGGCGAGTCGCAATAC 369
QY      225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
DB      370 TTTAATCAGCGCAGACTCAACACGAGCGAGTCCGGCTGAAAGAAAGAAATACCTCGGGCGT 429
QY      237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
DB      430 GGCAAACTCAATATCCAAATCACGCCCCAAAGTAAACCAACTCGCCGCCAACCGCGTGCAG 489
QY      250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
DB      490 ATCGACATTCACAGATTCAGAGGGCGAAATCCGCCAAATACCCGACATCGAATTTGAA--- 546
QY      270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289
DB      547 GGCACCAACGTCATTCATTCGACCGCCGCAAACTGATCGC-----CAG 585
QY      290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisHis 309
DB      586 ATGTCCTGACACAGGCGCGCATTTGGACATGGCTGCACAGCAAGCGACCGGTTGACCGC 645
QY      310 GlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPhe 329
DB      646 CAGAAATTCGCCCAAGACATGAAAGAAAGTAAACCGACTTCTACAGAAACAAAGCGCTACTTC 705
QY      330 AspGlyArgThrPheAsnAspArgSerValAspValIleLeuProAsnAsnThrAlaAspVal 349
DB      706 GATTTCGCTATTCCTCGATACCGACATCCAAACCAAGCAACAAACCAAGCAGACCATC 765
QY      350 SerLeuIleLysPheThrGlyThrGlnTyrArgPheAspGluValAlaPheIleThrIle 369
DB      766 AAATACCGCTCCACAGAAAGGCGGAGTTTCCGCTGGCGCAAGT-----TCGATT 816
QY      370 AspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeu 389
DB      817 GAAGGGGACACCAACGAA-----GTCCCAAGGCCGCA 849
QY      390 LeuGluGlnLeuLeuThrValAlaAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAla 409
DB      850 CTGGAATAACTGCTGCTGACCATGAAGCCCGCAAAATGTATGCAAGCCGACGATGACCGCC 909
QY      410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhe 429
DB      910 GTTTGGGT----- 918
QY      430 ProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArg 449
DB      919 -----GAGATTCGAACCGCATGGCTCGCCAGCAGTACGATACAC--- 960
QY      450 ThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeu 469
DB      961 -----GAAATCAGCGTACAGCGCTGCGCGAAGCGGAAACCGGAAACCA 999
QY      470 ThrAspGlyIleLeuMetAspIleSerPro-----IleGlu 481
DB      1000 ACCGCGATTTCTCTCTGCGCATCGAAGCGGCGGCAAAATTCACGTCAACGAAATTCAC 1059
QY      482 PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHis 501
DB      1060 ATCAGCGGCAACAAACAAACCGCGAGCAAA-----GTGTCGGCGCGGCAATTCGCCCA 1113
QY      502 LeuTyrAspMetProAspAsp-----ArgValIleuAla 512
DB      1114 ATGGAATTCGCGGCTTACACACACCTCCAAAGCTGCAACGCTCCAAAGAGCGCGTGCAGCTT 1173
QY      513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532
DB      1174 TTGGGCTACTTCGACAAACGTACAG-----TTTGAT 1203

```

```

QY 533 AlAValAlaArGAlaIleLeuProAspLysSerGluAsnGluValIleAspLeuProGlu 552
    |||||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1204 GCGCTCCCGCTGGCGGAGACCGCCGACAAA-----GTGATTTG----- 1242
QY 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValPro 572
    :|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1243 AACTGAGCCCTGACCGAAGCTCCAC----- 1269
QY 573 LeuYrValPheValAlaSerAspLysProArgAspGlnIleGlyLeu-----Gly 590
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1270 -----GGCTGCTGCTGACTTGACCGCGGC 1293
QY 591 TrpLysSerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuLeuAsnArg 610
    |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1294 TGGTTACAGATACCGG-----TTGGTCATGTCGGCGCGCGTATCCAGACAAACCTG 1347
QY 611 AspGlyTyrGlnAlaGlyAlaGlnLeuArgLeuSerGluAspLysLysGlyValLysLeu 630
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1348 TTCGGTACGGCAAGTCGGCGCGCTCGCGCTCGCAAGCAAA-----ACCAAGCTC 1401
QY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
    :|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1402 AACGGCTGCTGCTGTTTACCGACCCGTACTTCACGGCAAGCGGCGTGGGCTAC 1461
QY 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670
    :|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1462 -----GATATTACGGAAGACCTTCGACCGCGCAAGACATGACACGCTCAACAA 1515
QY 671 HisGluLeuSerArgSerIleIleGlnAsnGlyTyrPasnArgThrTyrSerLeuArg 690
    :|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1516 TATTAACCCACCCACCGC-----GGCGCGCGCTGATGATGGG-TATCCCGCT 1562
QY 691 TyrArgLeu-AspLysLeuLysThrGlnAlaProProGlu-----ThrTr 705
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1563 TACCGAATACGACCGCGCTCAATTGCGGCTGGCGCGCAACCTGACCGCTCAACACCTA 1622
QY 705 PglAspLeuPro-----ValAspPheValAsn-----GlyLys---ProSerG 719
    |||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1623 CAACAAGACACCAACCGCTATGCGGACTTATCAGAAATACGCAAAACGACGAGCGC 1682
QY 719 nGluAlaLeuLeuAlaGlyValAlaValHisLysThrVal-----Al 733
    :|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1683 AGACGGACGCTTCAAGGCGCTGTACAAAGCGCCTCGCGCGCGCGCAACAGAC 1742
QY 733 AspAsnLeuValAsnProMetArgGlyTyrArgGlnArgLysSerLeuGluValGlySe 753
    |||||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1743 CGACAGCGCGCTCATGGCGGCGCGGCTAC----- 1773
QY 753 rSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyLysSerGlyValTyr 773
    :|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1774 -----CTGACCGCGCTAAATGCGGAATGCGCCTGCGCGCGCAACAATCCTA 1823
QY 773 rSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGly----- 790
    :|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1824 C-----TACTCGCCACCCCAACCAACCACTGCTCTCTCC 1859
QY 791 -----GlyIleGlnAlaGlyTyrIle 797
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1860 CTTTAAGCAAAACCTTACCGCTGATGCTGGCGCGCAACTGCGCATTCGGCGGCTAC-- 1917
QY 797 eTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyLysAsp 817
    :|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1918 -----GGCAGAACCAAGAAATCCCTTCTTGAATACTTCAAGCGGCGGCTGG 1970
QY 817 nSerIleArgGlyTyrAlaHisAspSerLeuSerProLysSerAspLysGlyTyrLeuTh 837
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 1971 TTCGGTGGCGGCTACGAAACCGGACCGCTCGCGCG-----AAAGTGTATGACGA 2021
QY 837 rGlyGlyGlnValLeuAlaValGly-----ThrAlaGluTyrAsnTyrGluPhe 854
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2022 ATTAGCGGAAAAATACGCTACGGCGGCGCAACAAAAAACCAACGTCCTCGCGGCTGCT 2081

```

```

QY 854 t-----LysAspLeuArgLeuAlaValPheGlyAspIle 865
    :|||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2082 CTTCCCGATGCCCGGCGCAAGACCGACGACCGCTCCGCTGAGCTGTTCCACACG 2141
QY 865 eGly----- 866
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 2142 AGCAGCGCTGTGGACGCGAAGCTATACCGCGCGCAACGTAACAAATCGGT 2201
QY 867 -----AsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGly 883
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2202 TTACTCGGAAACCGCCATTAATCCACCTTTACCAAGAAATGCGCTATTCCCGCGCG 2261
QY 883 lGlyValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThrGlyVal 903
    |||  |||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 2262 CGCGGTACTGCTGCTGCTGCTTTGGTCCGATGAATTCAGTACGCTACCGCTGA 2321
QY 903 sGlu-----GluGlyAsnProIleLysLeuHisPheIleGlyThrProPhe 919
    :|||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 2322 GAAAAAACCGAAGCAAGAAATCAACGCTTCCATTCAGCTCGCGACGACGTTG 2376
RESULT 10
AAH42129
ID AAH42129 standard; DNA; 2379 BP.
AC AAH42129;
XX
DF 17-SEP-2001 (first entry)
XX
DE Nucleotide sequence of a Neisseria gonorrhoeae protein.
XX
KW Serogroup B protein; outer membrane protein; Neisserial infection;
KM vaccine; ss.
XX
OS Neisseria gonorrhoeae.
XX
FH Key 1.2379 Location/Qualifiers
FT CDS 1..2379
FT FT /*tag= a
FT FT /product= "Neisseria serogroup B protein"
FT sig_peptide 1..63
FT FT /*tag= b
FT FT mat_peptide 64..2376
FT FT /*tag= c
XX
PN MO200152885-A1.
XX
PD 26-JUL-2001.
XX
PF 17-JAN-2001; 2001WO-1B00166.
XX
PR 17-JAN-2000; 2000GB-0001067.
XX
PR 09-MAR-2000; 2000GB-0005699.
XX
PA (CHIR-) CHIRON SPA.
XX
PI Pizza M, Rappuoli R, Giuliani M;
XX
DR WPI: 2001-451895/48.
XX
DR P-PSDB: AAB84745.
XX
PT Composition for treating or preventing infection to, detecting, or for
PT raising antibodies against Neisserial bacteria, comprises an N.
PT meningitidis serogroup B outer membrane preparation and an immunogenic
PT component -
XX
PS Disclosure: Page 64-65; 83pp; English.
XX
CC The present sequence encodes a Neisseria gonorrhoeae protein. The
CC protein is used to produce the compositions of the invention. The
CC specification describes a composition, comprising a Neisseria
CC meningitidis serogroup B outer membrane preparation and an immunogenic
CC component. The immunogenic component is protein disclosed in WO99/57280,
CC WO99/36544, WO99/24578, WO99/66791, WO97/28273, WO96/29412, WO95/03413,

```

CC M099/21132, M099/58683, M099/55873, and/or N. meningitidis protein PorA, CC TdpA, TdpB, PilC, Opa, or Omp5. The composition is used for making CC a medicament for treating or preventing infection due to Neisseria CC bacteria; a diagnostic reagent for detecting the presence of Neisseria CC bacteria or of antibodies raised against Neisseria bacteria; and/or CC a reagent which can raise antibodies against Neisseria bacteria. It may CC also be used as a vaccine.

XX Sequence 2379 BP: 636 A; 755 C; 596 G; 392 T; 0 other;

Alignment Scores:

Score:	6.03e-12	Length:	2379
Percent Similarity:	265.50	Matches:	195
Best Local Similarity:	34.28%	Conservative:	120
Query Match:	21.22%	Mismatches:	348
	5.62%	Indels:	257
		Gaps:	39

US-09-914-168-2 (1-919) x AAH42129 (1-2379)

```
QY 116 ThrProLeuSerLeuGluGluIleuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
   ::::::::::::::::::::
DB 46 TCGCCTTTCGCATTTGCCACTTCCACCATCCGATCCGTCGAAAGCTTCACAGCT 105
   ::::::::::::::::::::

QY 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149
   ::::::::::::::::::::
DB 106 ACCGAGCCGACACCGATTATTCACACTGCGCTCAAGTCGCGACCTTACAC--- 162
   ::::::::::::::::::::

QY 150 GluValValValProProThrIleuGluProGluIleuIleuIleuIleuIleuIleu 169
   ::::::::::::::::::::
DB 163 -----GACACACACCGCAGCTGCCATCATCAAAAGCTGTAC 198
   ::::::::::::::::::::

QY 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuValLys 186
   ::::::::::::::::::::
DB 199 GCCACCGGTTTCTTTGACGACGAGTCGAAACTCGGACGCGGACGCTTGTGCTGACC 258
   ::::::::::::::::::::

QY 187 PheTyrClnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
   ::::::::::::::::::::
DB 259 GTATATGACGACCGCCACCATCGGCTCGCTCAACATCAGCGCGCAAAATGCTGCACAAAC 318
   ::::::::::::::::::::

QY 207 GluProTyrAlaAsnIleLysAlaIleuGluAsp-----IleThrGlnGluSerAla 224
   ::::::::::::::::::::
DB 319 GAGGCC-----ATCAAGAAACCTCGAATCGTTCGCGTGGCGCAGTCGCACATAC 369
   ::::::::::::::::::::

QY 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
   ::::::::::::::::::::
DB 370 TTTAATTCAGCGCACACTCAACACGAGCATGCGCGCTGAAGAAGAAATACCTCGGCGCT 429
   ::::::::::::::::::::

QY 237 -----ThrAlaLeuValAlaIleArgAlaValGlyTyrTyrAsp 249
   ::::::::::::::::::::
DB 430 GCGCAATCAATATTCCAATATCACGCCCAAGTAACCAACTCGCGCGCAACCGCGTGCAC 489
   ::::::::::::::::::::

QY 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
   ::::::::::::::::::::
DB 490 ATTCGACATCAGCATGTTGACGAGGCGCAATCGCGCAAAATCAACCGCATCGAATTGAA--- 546
   ::::::::::::::::::::

QY 270 GlyGluProValTyrIleAspTyrArgArgAlaValGluValArgIleGluIleValAspAsp 289
   ::::::::::::::::::::
DB 547 GCGAACCAATCTATTTCCGACCGCCCAACATGATCGG-----CAG 585
   ::::::::::::::::::::

QY 290 LysAlaPheThrThrValAlaAspGluValProLeuIleuIleGlyAspValPheHisHis 309
   ::::::::::::::::::::
DB 586 ATGTGCTGACCGAAGCGCGCATTTGACATGCGTTCACACGAAAGCAACCGGTTGACCGCGC 645
   ::::::::::::::::::::

QY 310 GlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPhe 329
   ::::::::::::::::::::
DB 646 CAGAAATTTGCGCCAAAGACATGCAAAAATACCGACTTCTACAGAAACGCGCTACTTC 705
   ::::::::::::::::::::

QY 330 AspGlyArgTyrPheLysAspArgSerValAspValIleLeuProAspAsnThrAlaAspVal 349
   ::::::::::::::::::::
DB 706 GATTTTCCGATATCTCGATACCGACATCCAAACCAAGCAAAACCAAGCAGACATC 765
   ::::::::::::::::::::

QY 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIle 369
   ::::::::::::::::::::
```

```
DB 766 AAATACACCTTCACGAAAGCGGAGCTTTCGCTGGCGCAAGTG-----TCGATT 816
   ::::::::::::::::::::
QY 370 AspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeu 389
   ::::::::::::::::::::
DB 817 GAAGCGCAGCCCAACGAA-----GTCCCAAGGCCCA 849
   ::::::::::::::::::::

QY 390 LeuGluGlnIleuLeuThrValAsnMetGlyAlaTyrAsnLeuGlnAlaValArgAla 409
   ::::::::::::::::::::
DB 850 CTGGAAATCTGTGACCATGACCAAGCCGCAAAATGCTACGAACCGCACAGATGACCGCC 909
   ::::::::::::::::::::

QY 410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhe 429
   ::::::::::::::::::::
DB 910 GTTTTGGCT----- 918
   ::::::::::::::::::::

QY 430 ProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArg 449
   ::::::::::::::::::::
DB 919 -----GAGATTCAAGACCGCATGGGCTGCGCAGCTACGCATACAGC----- 960
   ::::::::::::::::::::

QY 450 ThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeu 469
   ::::::::::::::::::::
DB 961 -----GAATACAGCGTACAGCGCTGCGCAACCGCAACCGCAACCA 999
   ::::::::::::::::::::

QY 470 ThrAspGlyIleLeuMetAspIleSerPro-----IleGlu 481
   ::::::::::::::::::::
DB 1000 ACCGTGATTTCTGCTGTCGACATCGAACCGCGCGCGCAAAATCTACGTCAACGAATATCCAC 1059
   ::::::::::::::::::::

QY 482 PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaLysAlaArgHis 501
   ::::::::::::::::::::
DB 1060 ATCAGCGGCAACACAAACCCGCGCA-----GTGCTGCGCGCGCAATTCGCGCA 1113
   ::::::::::::::::::::

QY 502 LeuTyrAspMetProAspAsp-----ArgValLeuAla 512
   ::::::::::::::::::::
DB 1114 ATGGAATCCGCGGCTTACGACACCTCCACAGCTGCACCGCTCCAAAGCGCGCTGACCTT 1173
   ::::::::::::::::::::

QY 513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532
   ::::::::::::::::::::
DB 1174 TTGGGCTACTTTCGACACGTAACAG-----TTTGAT 1203
   ::::::::::::::::::::

QY 533 AlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
   ::::::::::::::::::::
DB 1204 GCGCTCCCGGCTTCCGCGGTACCGCCGACAA-----GTGCAATTG----- 1242
   ::::::::::::::::::::

QY 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValPro 572
   ::::::::::::::::::::
DB 1243 AACATGAGCTTACCGCAACCTCCAC----- 1269
   ::::::::::::::::::::

QY 573 LeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeu-----Gly 590
   ::::::::::::::::::::
DB 1270 -----GGCTCGCTCGACTTGAGCGCGGCG 1293
   ::::::::::::::::::::

QY 591 TyrGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArg 610
   ::::::::::::::::::::
DB 1294 TGGGTTCCAGATATCCGCG-----TTGCTATGTCGCGCGCGCTATGCGAGCAACCTG 1347
   ::::::::::::::::::::

QY 611 AspGlyTyrGlnAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeu 630
   ::::::::::::::::::::
DB 1348 TTGCGTACGGGCAAGTCGCGCGCTGCGCGCTCGGAAAGCAAA-----ACCAGGCTC 1401
   ::::::::::::::::::::

QY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
   ::::::::::::::::::::
DB 1402 AACGCTCGCTGCTGTTTACCGACCGCTACTTCACGCAACGCGGCTGACCGCTGCTAC 1461
   ::::::::::::::::::::

QY 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670
   ::::::::::::::::::::
DB 1462 -----GATATTACGGAAGAGCTTCGACCGCGCAAGCATCGACAGCTCAACAAACA 1515
   ::::::::::::::::::::

QY 671 HisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThrTyrSerLeuArg 690
   ::::::::::::::::::::
DB 1516 TATAAACCAACCAACCGC-----GGCGGCGCGTAAAGATGGG-TATCCCGCT 1562
   ::::::::::::::::::::

QY 691 TyrArgLeu-AspLysLeuLysThrGlnAlaProProGlu-----ThrTr 705
   ::::::::::::::::::::
```



```

Db 1563 TACCGAATACGACCGCTCATTTTCGGGCTGGCGGACACCTGACCGTCAACACCTA 1622
QY 705 pGlnAspLeuPro-----ValAspPheValAsn-----GlyLys---ProSerGI 719
Db 1623 CAACAAAGCACCCAAAGCCTATGCCGACTTATATCAGGAATACGCAAAACGACGCGGC 1682
QY 719 ngUlaLeuLeuAlaGlyValAlaValHisLysThrVal-----Al 733
Db 1683 AGACGGCAGCTTCAAGAGCCCTGCTGACAAAGCAGCCGCTGGGGGCGCAACAAAGAC 1742
QY 733 aAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySe 753
Db 1743 CGACAGCGCGCTCATGCGGACGCGGTAC----- 1773
QY 753 rSerGlyLeuValSerAspAlaAsnMetAlaLeuAlaArgAlaGlyLysSerGlyValTyr 773
Db 1774 -----CTGACCGCGCGTAAATGCCGAATGCGCCCTGCCGCGACCAACATGCAATA 1823
QY 773 rSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGly----- 790
Db 1824 C-----TACTCCGCCACCAACCAACCAACCTGCTCTCC 1859
QY 791 -----GlyLysGlnAlaGlyTyrI 797
Db 1860 CTTAAGCAAAACCTTACAGCTGATGCTGGCGGCAAGCTGCGCATTCGGCGGCGCTAC-- 1917
QY 797 eTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyLysAspGI 817
Db 1918 -----GGCAGAACCAAGAAATCCCTTCTTGAAAACTTCAAGCGCGCGCTGSGG 1970
QY 817 nSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuTh 837
Db 1971 TTCGTGCGCGGCTACGAAACGCGACGCTGCGCCG-----AAAGTGATGACGA 2021
QY 837 rGlyGlyGlnValLeuAlaValGly-----ThrAlaGluTyrAsnTyrGluPheMe 854
Db 2022 ATACGGCGAAATAATCAAGCTACGCGGCAACAAACCAACGCTCTCGCGGAGCTGCT 2081
QY 854 t-----LysAspLeuArgLeuAlaValPheGlyAspII 865
Db 2082 CTTCGCCATGCGCGCTGCGAAAGACGACGACGCTGCGCTGAGCCCTGTTGCCGACGC 2141
QY 865 eGly----- 866
Db 2142 AGCAGCGCTGCGGACGACGACCTATACGCGCGCGAAACGTTACACAAATCGGT 2201
QY 867 -----AsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyVa 883
Db 2202 TTAAGTGGAAACCGCATAAATCCACCTTTACACCAATTTGCGGTATTCGCGCGGG 2261
QY 883 IglYValArgTyrAlaSerProValIglYValArgValAspValAlaThrGlyValY 903
Db 2262 CGCGGTTACCTGCGCTCGCTTTGGGTCCGATGAATTCAGCTACGCTTACCGGCTAA 2321
QY 903 sGlu-----GluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
Db 2322 GAAATAAACCGAGACAGCAATCATCAAGCTTCCAAATTCAGCTGCGGACAGAGTTTC 2376

```

RESULT 11

AAAI5155 standard; DNA; 2505 BP.

AAAI5155;

21-AUG-2000 (first entry)

DNA encoding outer membrane protein (omp) 85.

Outer membrane protein; omp85; gonococcal infection; symptomatic disease;

meningococcal infection; protective immune response; vaccine; ss.

Neisseria gonorrhoeae.

XX

```

FH Key Location/Qualifiers
FT RBS 49..53
FT 5'UTR 1..59
FT CDS 60..2438
FT 3'UTR 2439
FT sig_peptide "outer membrane protein 85"
FT 60..122
FT 2439
FT /*tag= d
FT /*tag= e
W0200023595-A1.
27-APR-2000.
22-OCT-1998; 98WO-US22352.
22-OCT-1998; 98WO-US22352.
(UYMO-) UNIV MONTANA.
JuaD RC, Manning SD;
WPI: 2000-339694/29.
DR P-PSDB; MAY84946.
XX
XX Claim 6; Fig 2; 98pp; English.
XX
XX The present sequence encodes an outer membrane protein (omp) 85 of
CC Neisseria gonorrhoeae. The omp polypeptides and polynucleotides are
CC useful in compositions for use in the prevention, treatment and diagnosis
CC of non-symptomatic gonococcal infection or meningococcal infection and
CC symptomatic disease. They are also useful for the detection of
CC hybridisation complexes. Antigens and antibodies specific omp proteins
CC also provide diagnostic, therapeutic and prophylactic compositions for
CC the treatment or prevention of the infections described above. The
CC antibodies are useful for inducing a protective immune response in
CC humans or animals with N. gonorrhoeae, N. meningitidis, or other
CC Neisseria species. The proteins, antibodies and polynucleotide
CC sequences of the present invention may also be used in the screening
CC and development of chemical compounds such as drugs or vaccines.
XX
XX Sequence 2505 BP; 669 A; 787 C; 621 G; 428 T; 0 other;
XX
XX Alignment Scores:
Pred. No.: 1.99e-11 Length: 2505
Score: 259.50 Matches: 192
Percent Similarity: 33.95% Conservative: 120
Best Local Similarity: 20.89% Mismatches: 351
Query Match: 5.49% Indels: 257
DB: 21 Gaps: 37
US-09-914-168-2 (1-919) x AAAI5155 (1-2505)
QY 116 ThrProLeuSerLeuGluLeuPheAlaGlnGluSerThrGluMetGlyLysAsnPro 135
Db 105 TCGGCTTTGGCATTTGCCGACTTCACATCCAGACATCCGTCGGAAGGCTTGACAGGT 164
QY 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149
Db 165 ACCGAGCCGAGACCGCTATTCACACTACCTGCCGTCAAAGTCGCGACACCTTACAC-- 221
QY 150 GluValValAlaProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
Db 222 -----GACACACAGCGCAGTGCCTCATCATCAAAAGCCTGTAC 257
QY 170 AlaArgLeuPheAsnAspGlyVal----- 177

```

Db 258 GGCACCGGTTCTTTGACGACGTACGATCGAAGTCGCGACGGCTGCTTCTGCTGACC 317
178 -----AsnLysValProArgLeuLysAlaLysPheTyGlnSer 190
318 GTTATGCTATGCCCTACATCGGCTGCTCACATCACCGCGCCCAAAATGCTCGCAGAAC 377
191 SerGln---SerGlyLthrSerAlaIleGlySerSerHisGlnLysThrGluProTyr 209
378 GACGCCATCAGAAAAAACCCTCGAATGCTTGCGGCTGCGGCTGCGAATCGCAATCTTAATCAG 437
210 AlaAsnIleLysAlaIleLeuGlnAspIleThrGlnGlnSerAlaMetLysPheLysGly 229
438 GCGACACTCACCGGACGTGCGCGGCTGAAAGAAATAT-----CTCGGGCGC 488
230 SerIleProArgLthrArgLthrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
489 GCGAAACTCAATATTCAAATACAGCCCAAGTACCAAACTGCCCGCAACCGCGTCGAC 548
250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
549 ATCGACATCACGATGACGAGCGCAAAATCCGCCAAATCACGACATCGAATTTGAA--- 605
270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289
606 GGCAAACCACTCTATTCCGACCGCAAACTGATGCGG-----CAG 644
290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisHis 309
645 ATGTGCTGACCGAAGCGGCAATTTGGACATGCTGACACGAAAGCAGCGGTTGACACCGC 704
310 GlyLysTyrGluThrLysLysAsnLeuIleGlnAsnAlaSerAlaGlnHisGlyTyrPhe 329
705 CAGAAATTCGCCCAAGACATGCAAAAGTACCGACTTACCGAAGAACGCGCTACTTC 764
330 AspGlyArgTyrPheAspArgSerValAspValIleLeuProAspAsnThrAlaAspVal 349
765 GATTTCCGATCTCTGATACCGACATCCAAACCAACGAAACCAACGACGACCATC 824
350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhePheThrIle 369
825 AAATTCACCGTCACGACGAGCGGACGTTCCGCTGGCGCAAACTG-----TCGATT 875
370 AspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeu 389
876 GAAGCGCACCAACGAA-----GTCCCAAGCGCGGAA 908
390 LeuGlnGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAla 409
909 CTGGAAGAACTGCTGACCATGAAAGCCCGCAATGCTACGAAACCGACGATGACCGCC 968
410 LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhe 429
969 GTTTTCGCT----- 977
430 ProGluArgGlnGlnIleGlnAsnAspGlnValSerPheGlnGlnSerSerSerArg 449
978 -----GAGATTCCAGAACCCCATGGCTCGGAGCTACCATACAC----- 1019
450 ThrGluProAlaGlnValAspGlnSerThrLeuGluProValIleGluThrValGluLeu 469
1020 -----GAAATCAGCGTACAGCCGCTGCCGAACCGCGAACCAAA 1058
470 ThrAspGlyIleLeuMetAspLysPro-----IleGlu 481
1059 ACCGTGATTTGCTGCTGACATCGAAGCGGAGAAAATCTACGTCAAGAAATTCAC 1118
482 PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaIleLysAlaArgHis 501
1119 ATACCGCGCAACAAACCAACCGCGACGAA-----GTGCTGCGCGCGAATTCGCCCA 1172
502 LeuTyrAspMetProAspAsp-----ArgValLeuAla 512
::: 111 111

Db 1173 ATGAATCCGGCGCTTACGACACCTCCAACTGCAACGCTCCAAAGACGCGTGACCTT 1232
513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAlaValSer 532
1233 TTGGGCTACTTCGACAACGTCACAG-----TTTGAT 1262
533 AlaValAlaArgAlaIleLeuProAspGlnSerGlnAsnGluValIleAspLeuProGlu 552
1263 GCCCTCCCGCTTCGCGGTACGCGCAAA-----GTGCTTTG----- 1301
553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValPro 572
1302 AACATGAGCTTACCGAAGCTTCAC----- 1328
573 LeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeu-----Gly 590
1329 -----GGCTCCGCTGACCTTGACCGCGC 1352
591 TrpGlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArg 610
1353 TTGGTTTCAGGATACCGC-----TTGGTCATGTCGCGCGCTATCGACGACAACTCG 1406
611 AspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGlnLysLysGlyValLysLeu 630
1407 TTGCGTACGGCGCAAGTCGCGCGCGCTGCGGCTCGCGAAGCAAA-----ACCAGCTC 1460
631 TyrAlaThrLysProLeuSerHisProLeuAsnAspLeuValArgAlaThrLeuGlyTyr 650
1461 AACGGCTCGCTGCTGTTTACCGACCGCTACTTACGCGCAACGCGGCTGACCTGGGCTTC 1520
651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670
1521 -----GATTTTACGGAAGAGCTTCGACCGCGCAACATCGACACGCTCAACAA 1574
671 HisGlnLysSerArgSerIleIleGlnAsnGlyTyrPheAsnArgThrTyrSerLeuArg 690
1575 TATMAAACCAACACCGC-----GGCGGCGGCTAAGATGGG-TATCCCGT 1621
691 TyrArgLeuAspLysLeuLysThrGlnAlaProGlu-----ThrTr 705
1622 TACCGAATACGCGCGGTCAATTCGGGCTGGCGGCGAACACTGACCTCAACACTTA 1681
705 PheAspLeuPro-----ValAspPheValAsn-----GlyLys---ProSerG 719
1682 CAACAAAGCACCAACGCTATGCGCATTCATCAAACTACGCGCAACGACGCGCGC 1741
719 ngluAlaLeuLeuAlaGlyValAlaValHisLysThrVal-----Al 733
1742 AGACGGGAGCTTCAAGCGCTGCTGTCACAAAGCACACTGTGCGGTGGGCGCAACAGAC 1801
733 aspAsnLeuValAsnProMetArgLysTyrArgGlnArgTyrSerLeuGluValGlySe 753
1802 CGACAGCGCTTATGGCGGACGCGCGCTAC----- 1832
753 rSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyr 773
1833 -----CTGACCGGCTTAATGCGGAATTCGCGCTCCGCGCGCAACTGCAATA 1882
773 rSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGly----- 790
1883 C-----TACTCCGCGCACCCACACCAACCAACTGGTTCTTCCC 1918
791 -----GlyIleGlnAlaGlyTyrIle 797
1919 CTTAAGCAAACTTCACCGCTGATGCTCGCGGCGAAGTGGGATTCGGGCGGCTAC-- 1976
797 eTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyLysAspG 817
1977 -----GCGAGAACCAAGAAATCCCTCTTTGAAATCTGTACGCGCGCGCTGGG 2029
817 nSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuTh 837
2030 TTGCGTGGCGGCTTACGAAAGCGGACGCTGCGCGC-----AAAGTATGACGA 2080

QY 837 rGlyGlnValLeuAlaValGly-----ThrAlaGluTyrAsnTyrGluPheMe 854
 D 2081 ATACGGCGAAAAATCAGTCAGCGGCAACAAAAAGCCAAACGCTCTCCGCGAGCTGCT 2140
 QY 854 t-----LysAspLeuArgLeuAlaValPheGlyAspII 865
 D 2141 CTTCCGATGCCGGTGGAAAGACGACGACCGCTCCGCTGAGCCTGTTGGCGAGC 2200
 QY 865 eGly----- 866
 D 2201 AGGACGCTGTGGACGACCACTATACCGCGCCGAAACGATACACAATCCGCT 2260
 QY 867 -----AsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyVa 883
 D 2261 TTACTCGGAAACGCGATAAATCCACCTTTTACCAACGAATTCGCTATTCGCGGCGG 2320
 QY 883 lGlyValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThrGlyVally 903
 D 2321 CGCGGTTACCTGGCTCTCGCCTTTGGGCCGATGAAATTCATCTACGCTACCGCTGAA 2380
 QY 903 sGlu-----GluGlyAsnProIleLysLeuHisPheIleGlyThrProPhe 919
 D 2381 GAAAAACCGAAGACGAAATCCACGCTTCCAAATTCAGCTCGCGACGCTTC 2435
 RESULT 12
 AAS07277
 ID AAS07277 standard; DNA; 2394 BP.
 XX
 AC AAS07277;
 DT 23-OCT-2001 (first entry)
 XX
 DE Neisseria meningitidis serogroup B antigenic protein DNA.
 XX
 KW Serogroup B antigen; pharynx; meningitis; septicaemia; mammalian cell;
 KW bacterial infection; baculovirus; yeast; ds.
 XX
 OS Neisseria meningitidis.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2394
 FT /tag= a
 FT /product= "N. meningitidis serogroup B antigen"
 FT sig_peptide 1..63
 FT /tag= b
 FT mat_peptide 64..2391
 FT /tag= c
 FT /product= "Mature N. meningitidis serogroup B antigen"
 XX
 PN W0200138350-A2.
 XX
 PD 31-MAY-2001.
 XX
 PF 28-NOV-2000; 2000WO-IB01851.
 XX
 PR 29-NOV-1999; 99GB-0028197.
 PR 09-MAR-2000; 2000GB-0005698.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (STAT-) STATENS INST FOLKEHELSE.
 XX
 PI Giuliani MM, Pizza M, Rappuoli R, Holst J;
 XX
 DR WPI: 2001-381289/40.
 DR P-PSDB: AAU03957.
 XX
 PT Novel 85 kda antigen from Neisseria meningitidis and Neisseria
 PT gonorrhoeae, useful in the manufacture of a medicament for treating and
 PT preventing Neisserial bacteria infection -
 XX
 PS Claim 6; Fig 2; 92pp; English.
 XX

CC The sequence represents a DNA encoding a Neisseria meningitidis serogroup
 CC B 85 kda antigenic protein. Neisseria meningitidis colonises the pharynx,
 CC causing meningitis and, occasionally, septicaemia in the absence of a
 CC meningitis. This antigenic protein is useful in the manufacture of a
 CC medicament for treating or preventing infection due to Neisseria
 CC bacteria, such as meningitis and septicaemia. It is also useful as a
 CC diagnostic reagent for detecting the presence of Neisseria bacteria or
 CC antibodies raised against Neisseria, and as a reagent for raising the
 CC antibodies. The Neisserial nucleotide sequences can be expressed in a
 CC variety of different expression systems, for example, mammalian cells,
 CC baculoviruses, plants, bacteria and yeast.
 XX
 SQ Sequence 2394 BP; 657 A; 743 C; 583 G; 411 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,06e-11 Length: 2394
 Score: 259.00 Matches: 191
 Percent Similarity: 33.48% Conservative: 121
 Best Local Similarity: 20.49% Mismatches: 342
 Query Match: 5.48% Indels: 278
 DB: 22 Gaps: 40
 US-09-914-168-2 (1-919) x AAS07277 (1-2394)
 QY 116 ThrProLeuSerLeuGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
 D 46 TCGCCTTTGGCACTTCCGACTTCACCATCCAAAGACATCCGCGTCGAAGCTTGCAGCGT 105
 QY 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149
 D 106 ACCGAGCGGAGTACCGTATTCAACTACCTGCCCGCTCAAAAGTCGGCGACACCTACAAC-- 162
 QY 150 GluValValProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
 D 163 -----GACACACACGCGAGTGCCTATCATCAAAACCTGTGTAC 198
 QY 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
 D 199 GCCACCGGTTCTTTGACGACGTACGCGTCGAAACTCGCGAGCGGCGAGCTCTGTGTACC 258
 QY 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
 D 259 GTTATCGAACGCCGCCACCATCGCTCGCTCAACATCATCCGCGGCAAAATGCTGCAAAAC 318
 QY 207 GluProTyrAlaAsnIleLysAlaAlaLeuGluAsp-----IleThrGlnGluSerAla 224
 D 319 GACGCC-----ATTAAGAAACCTCGAATCGTTCGGGCTGGCGAGTGCATATC 369
 QY 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
 D 370 TTTAATCAGCGGACACTCAATCAGCGAGTCCGCGGCTGAAAGAGATAATACCTCGGCGCG 429
 QY 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
 D 430 GGCAAACTCAATATCAATATCCCAAGTAAACCAAACTCGCCCGCAACCGCGTCGAC 489
 QY 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
 D 490 ATCGACATCAGATTGACGAGGGCAAAATCCGCCAAATCACCACATCAATTTGAA--- 546
 QY 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289
 D 547 GGCAACCAAGTCTATTCCGACCGCAAACTGTATGCGG----- 582
 QY 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
 D 583 -----CAAAATGTCTCTGACCGCAAGCGCGCATTTGGACATGG 618
 QY 307 -----PheHisGlyLysTyrGluThrLysLysAsnLeuIleGlu 320
 D 619 CTGACACCAAGCAACCAATTCACGAGCAGAAATTTGCCCAAGATATGAAAAAGTAACC 678
 QY 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAspArgSerValAspVal 340

Db	679	GA	CTTCTACCAAAATACCGGCTACTTCGATTTCCGATTCCTCGATACCGACATCCAAAC	738
Qy	341	Ile	LeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArg	360
Db	739	AC	GAAGACAAAACCAAGCAGACCATCAAAATCACCGTCCACGAAGCGGACGTTCCCGT	798
Qy	361	Phe	AspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro	380
Db	799	TGGG	CAAGTC-----TCCATCGAAGCGACACCAACGAA-----	834
Qy	381	Asp	LysLeuProValLysArgGluLeuLeuLeuGluLeuLeuThrValAsnMetGlyGlu	400
Db	835	-----	GTCCCAAGCCGAACCTGGAAAACTGCTGACCATGAAGCCGCGCAAA	882
Qy	401	Ala	TyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe	420
Db	883	TGGT	ACGACGCCAGCATGCCCGCTTTGGGT-----	918
Qy	421	Asn	MetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnVal	440
Db	919	-----	-----GAGATTTCAGAACCGCATGGCG	939
Qy	441	Ser	PheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu	460
Db	940	TCGG	CAGGCTACGCATACAGC-----	972
Qy	461	Glu	ProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro--	479
Db	973	CAGC	CGCTCCGCAACCTGAAACCAACCGTCGATTTCTGCTCGCATCGAACCGGGC	1032
Qy	480	-----	-----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu	492
Db	1033	CGGAAA	TTAGTCAACGAAATACATACCGCGCAACAACAAACCGCGACGAA--	1089
Qy	493	Asn	LeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspasp-----	508
Db	1090	----	GTGCTCGCGGTGAATTACGCCAAATGGAAATCCGCACCTTACGACACCTCCAAGCTG	1146
Qy	509	-----	-----ArgValLeuAlaIleAsnHisAspGlyValAsnArgSerIle	523
Db	1147	CAAC	GTTCCAAAGACGCGCTCGAGCTTTGGGCTACTTCGACAATGTCCAG-----	1197
Qy	524	Leu	GlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer	543
Db	1198	-----	-----TTTGATGCTGCCGCTTCCGCGCACCGCCGACAAA--	1233
Qy	544	Glu	AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla	563
Db	1234	-----	-----GTGATTTG-----AACATGAGTCTACCGAAGCTTCCACC-----	1269
Qy	564	Asp	ValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg	583
Db	1269	-----	-----	1269
Qy	584	Asp	GlyGlnIleGlyLeu-----GlyTyrPglySerAspThrGlyThrArgLeuValThr	601
Db	1270	--	GGTTCCTGGATTGAGCGGGTTCGGTTCAGATACCGGG-----TTGGTTCATG	1320
Qy	602	Lys	PheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu	621
Db	1321	TCGC	CAGCGGTTTCCCAAGACACACCTGTTCGGTACGGGCAAGTTCGGCGCACCTCGCGCC	1380
Qy	622	Ser	GluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn	641
Db	1381	TCGAG	GAGCAAA-----ACCAGCGTTAACGGCTGCTGTCGTTCATACCGCCTACTTC	1434
Qy	642	Asp	GlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly	661
Db	1435	ACGG	CAGCGGGTTCAGCTGGGTAC-----GATGTTTCGGAAGAAAGCTTCGACCGC	1488
Qy	662	Phe	AspLeuSerThrThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly	681

Db	1489	CGCAAGCATCGACCGACATCAACAACATATAAAACCAACCAACGGCA-----GGC	1536
Qy	682	GlyTrpAsnArg-----ThrTrpSerLeuArgTyrArgLeuAspLysLeu	696
Db	1537	GCAGGCATCCGCATGACGCGTGCCTGTTACCGAATACGACCGCTGAATTCGGTTGGTG	1596
Qy	697	LysThrGlnAlaProProGluThrTrpGlnAspLeuPro-----ValAspPheVal	713
Db	1597	GCAGAACACCTGACCGGTCAACACCTACAACAAGCGCCCAACACTATGCGCACTTATC	1656
Qy	714	Asn-----GlyLys---ProSerGlnGlnAlaLeuLeuAlaGlyValAlaValHisLys	730
Db	1657	AAGAAATACGGCAAAACCGACGCGACAGCGCAGCTTCAAGGCTGGCTGTACAAAGGT	1716
Qy	731	ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg	744
Db	1717	ACCTCGCTGGGGGGCAACAAACCGACAGCGGTATTGCCGACGCGCGGTAC---T	1773
Qy	745	GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle	764
Db	1774	-----CTCAGCGGCGTGAACGCGCGCAATC	1797
Qy	765	AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAlaTyrGlySerAsnArg	784
Db	1798	CCCTGCCTGGCAGCAACCTGCAATAC-----TACTCCGCC	1833
Qy	785	AlaHisGlnMetThrGly-----	790
Db	1834	ACCCCAACCAACACTGTTCTTCCCCCTGACGAAACCTTCACGCTGATGCTCGCGCGC	1893
Qy	791	-----GlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArg	808
Db	1894	GAAGTCGCATTCGGGCGGCTAC-----GGCAGAACCAAGAAATCCCTTCCTTT	1944
Qy	809	LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer	828
Db	1945	GAAGAACTTCTACGGCGCGCGCTGGTTCGGTCGCGGATACGAAAGCGGACGCTCGT	2004
Qy	829	ProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly-----	845
Db	2005	CCG-----AAAGTCTATGACGAATACGGCGGAAATAATCAGTACGGCGGCAACAA	2055
Qy	846	ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp	856
Db	2056	AAAGCCAAAGTCTCCGCGAGCTGCTTTCGCGATGCCGCGCGCAAGACGCGCGCAC	2115
Qy	857	LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp-----	870
Db	2116	GTCCGCTGAGCCTGTTTGGCGACGCGACGCGACGCGTGTGGGACGGCAAAACCTACGACGAC	2175
Qy	870	-----	870
Db	2176	AACAGCATTTCCGCGACGCGCGACGGGTTCAAAACATTTTACGGCGCGCGCAATACCCAT	2235
Qy	871	--LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer	889
Db	2236	AAATCCACCTTTTACCAACGAATTCGCGCTATTTCGCGCGCGCGGCTTACCTTGGCTCTCG	2295
Qy	890	ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GlyGlyAsn	907
Db	2296	CCTTAGCCCGATGAATTCACGTACGCTACCGCTACCCGCTGCAAGAAAAAACCGGAAGACGA	2355
Qy	908	ProIleLysLeuHisPheIleGlyThrProPhe	919
Db	2356	ATCCAACGGCTTCAATTTCCAACCTCGCAGCAGCGTTC	2391

RESULT 13

RESOL1 13
AAS07279
ID AAS07279 standard; DNA: 2394 BP.

AA
AC AAS07279:XX
XX
AC
AA301213;

DT 23-OCT-2001 (first entry)


```
Db 1033 CGGAAATCTACGTCAAGAAATCCACATCACCGGCAACAACAAACCCGGGAGGAA --- 1089
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 493 AsnLeuValAlaLysAlaArgHisLeuTyrAspMetProAspAsp- --- 508
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1090 ---GTCGGTGGCCGCGAATTCGGCAATGGAATCCCGCGCTTACGACACCTCCAAAGCTG 1146
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle 523
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1147 CAACGCTCCAAAGAGCGGTGAGCTTTGGGTACTTCGACAAAGCTACAG- --- 1197
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1198 -----TTTGATCCGCTCCGCTTGGCGGCACACCGCAACA --- 1233
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1234 -----GTCGATTTG-----AACATGAGCCTTGACCGAAGCTTCACCC- --- 1269
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1269 ----- 1269
Qy 584 AspGlyGlnIleGlyLeu-----GlyTrpGlySerAspThrGlyThrArgLeuValThr 601
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1270 ---GGCTCGCTCGACTTGAGCGCGGCTGGGTACAGGATACCGGC- ---CTGCTCATG 1320
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 602 LysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1321 TCCGAGCGGTTTCCCAAGACAACCTGTCGTACGCGCAAGTCGGCGCGCTGCGCGCC 1380
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1381 TCACGAAGCAAA-----ACCAGCTCAACGCTCGCTGCTTTACCGACCCGCTACTTC 1434
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1435 ACGCAGACGCGGTGACCTGGCTGCTAC-----GATGTTTACGGAAAGCCCTTCGACCCG 1488
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 662 PheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly 681
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1489 GCCAAGACATCGACCGAGCATCAACAATATAAACACCACCCAGCA-----GGC 1536
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 682 GlyTrpAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1537 GCAGGCATCCGCGATGAGCGTGCTGTACCGAATACGACCGCGTGAATTTCCGTTTGGTG 1596
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 697 LysThrGlnAlaProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1597 GCAGAACACCTGACCGTCAACACCTTACAAACAAAGCCCAACACACTATGCCGACTTTATC 1656
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1657 AAGAAATACGGCAAAACCGACGCGACAGCGCAGCTTCAAAGGCTGGCTGTACAAAGGT 1716
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1717 ACCGTGCGCTGGGGCGCAACAAACCGACAGCGGTTATGGCCGACGCGCGCTAC--- 1773
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 745 GlnArgTyrSerLeuGluValGlySerGlyLeuValSerAspAlaAsnMetalIle 764
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1774 -----CTGACGGGCGTGAACCCCGAAATC 1797
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1798 GCCCTGCCGCGCAGCAAACTGCAATAC-----TACTCCGCC 1833
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 785 AlaHisGlnMetThrGly----- 790
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1834 ACCCAACAAACCAACCTGGTTCTTCCCTTAAGCAAAACCTTCACGCTGATGCTCGGCGGC 1893
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 791 -----GlyIleGlnAlaGlyTyrIleIleTrpSerAspAsnHisValProTyrArg 808
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
```

```
Db 1894 GAAGTCGCATTCGGCGGCGGTAC-----GCAGAACCAAAAGAAATCCCTTCTTT 1944
Qy 809 LeuArgPhePheAlaGlyLysAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 1945 GAAAACTTCTACGGCGCGGCTGGGTTCGGTCCGGGATACGAAAGCGGCGCTCGGT 2004
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValIleGly----- 845
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 2005 CCG-----AAAGTGTATGACCAATACGGCGAAAAAATCAGCTACGGCGGCAACAA 2055
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 2056 AAAGCAACAGCTCCGCGGAGCTGCTTCCCGATGCCGCGCGGCGGCAACGCGGCACC 2115
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 2116 GTCCGCTGAGCCTGTTTCCCGACCGCAGCGCTGGGACGCGCAAAACCTACGACGAC 2175
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 870 ----- 870
Db 2176 AACAGCAGTTCCGCGACCGCGGAGGTTCAAAACATTTACGGCGCGGCAATACCCAT 2235
Qy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer 889
      |||      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 2236 AAATCCACCTTTACCAACAGATTCGCTATTCGCGCGCGCGGTTACCTGGCTCTCG 2295
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 2296 CTTTATAGCCGATGAAATTCAGCTACGCTACCGCTCAAGAAAAACCGGAAGACGAA 2355
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Qy 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
Db 2356 ATCCAACGCTTCCAATTCCAACTCCGACGCGACGAGCTTC 2391
      :|:      :|:      :|:      :|:      :|:      :|:      :|:      :|:
      RESULT 14
      AAH42128
      ID AAH42128 standard; DNA; 2394 BP.
      AC AAH42128;
      DT 17-SEP-2001 (first entry)
      DE Nucleotide sequence of a Neisseria serogroup B protein.
      KW Serogroup B protein; outer membrane protein; Neisserial infection;
      KW vaccine; ss.
      OS Neisseria meningitidis.
      FH Key Location/Qualifiers
      FT CDS 1..2394
      FT /*tag= a
      FT /product= "Neisseria serogroup B protein"
      FT sig_peptide 1..63
      FT /*tag= b
      FT mat_peptide 64..2391
      FT /*tag= c
      PN WO200152885-A1.
      XX 26-JUL-2001.
      PD
      XX 17-JAN-2001; 2001WO-IB00166.
      XX 17-JAN-2000; 2000GB-0001067.
      PR 09-MAR-2000; 2000GB-0005699.
      XX (CHIR-) CHIRON SPA.
      XX Pizza M, Rappuoli R, Giuliani M;
      XX WPI; 2001-451895/48.
      DR P-PSDB; AAB84744.
```

XX Composition for treating or preventing infection to, detecting, or for
 PT raising antibodies against Neisserial bacteria, comprises an N.
 PT meningitidis serogroup B outer membrane preparation and an immunogenic
 PT component

XX Disclosure: Page 58-59; 83pp; English.

FS The present sequence encodes a Neisseria serogroup B protein. The
 CC protein is used to produce the compositions of the invention. The
 CC specification describes a composition, comprising a Neisseria
 CC meningitidis serogroup B outer membrane preparation and an immunogenic
 CC component. The immunogenic component is protein disclosed in WO99/57280,
 CC WO99/36544, WO99/24578, WO99/66791, WO97/28273, WO96/29412, WO95/03413,
 CC WO99/31132, WO99/58683, WO99/55873, and/or N. meningitidis protein PorA,
 CC TbpA, TbpB, PilC, Opa, or Omp85. The composition is used for making
 CC a medicament for treating or preventing infection due to Neisserial
 CC bacteria; a diagnostic reagent for detecting the presence of Neisserial
 CC bacteria or of antibodies raised against Neisserial bacteria; and/or
 CC a reagent which can raise antibodies against Neisserial bacteria. It may
 CC also be used as a vaccine.

SQ Sequence 2394 BP; 657 A; 743 C; 583 G; 411 T; 0 other;

Alignment Scores:

Pred. No.:	2,06e-11	Length:	2394
Score:	259.00	Matches:	191
Percent Similarity:	33.48%	Conservative:	121
Best Local Similarity:	20.49%	Mismatches:	342
Query Match:	5.48%	Indels:	278
DB:	22	Gaps:	40

US-09-914-168-2 (1-919) x AAH42128 (1-2394)

Qy	116	ThrProLeuGluGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro	135
Db	46	TCGCCTTTGGCACTTCCCGACTTCACCATCCAGACATCCCGCTCGAAGGCTTGCGAGGT	105
Qy	136	AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer	149
Db	106	ACCGAGCGGAGTACCGTATTCAACTACCTGCCGCTCAAAATGCGGCGCACCTTACAAC---	162
Qy	150	GluValValProThrLeuProGluProGluLysProGlyLeuIleIysArgLeuTyr	169
Db	163	-----GACACACACGCGAGTGCATCATCAATAAAGCCTGTAC	198
Qy	170	AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys	186
Db	199	GCACCGGTTTCTTTGACGACGTACGCGTCAAACTCGGACGGCGAGCTCTCGTGACC	258
Qy	187	PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr	206
Db	259	GTATCGAAGCGCCCACTACCTCGCTCAACATCACCGCGCAAAATGCTGCACAAAC	318
Qy	207	GluProTyrAlaAsnIleLysAlaIleLeuGluAsp-----IleThrGlnGluSerAla	224
Db	319	GACGCC-----ATTAAAGAAACCTCGAATCGTTGCGGCTGGCGGCGAGTCGCAATAC	369
Qy	225	MetAsp-----LeuAsnGlySerIleProArgLeuArgGln-----	236
Db	370	TTTAAATCAGCGACACTCAATCAGGACGTCCGCGCGCTGAAGAGAAATACCTCGGGCGC	429
Qy	237	-----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp	249
Db	430	GCACAACTCAATATCAAAATCAGCCCAAGTAACCAAACTCGCCGCAACCGGCTCGAC	489
Qy	250	IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu	269
Db	490	ATCGACATCAGGATTCGACGAGGCAAAATCCGCCAAATCACCAGACATCGAATTTGAA---	546
Qy	270	GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluAlaAspAsp	289
Db	547	GGCAACCAAGTCTATTCCGACCGCAAACTGATGCGG-----	582

Qy	290	LysAlaPheThrValAlaAspGluValProLeuLeuIleGlyAspVal-----	306
Db	583	-----CAAAATGTCCCTGACCGAAGCGGCATTTGGACATGG	618
Qy	307	-----PheHisGlyLysTyrGluThrLysLysAsnLeuIleGlu	320
Db	619	CTGACACGAAGCAACCAATTCACGAGCAGAAATTTGCCCAAGATATGAAAGAAATGACC	678
Qy	321	AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTTPLeuAspArgSerValAspVal	340
Db	679	GACTTCTACCAAAATACGGCTACTTCGATTTCCGTATCTCTCATACCCACATCCAAACC	738
Qy	341	IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArg	360
Db	739	AACGAAGACAAACCAAGCAGACCATCAAAATCCCGTCCACGAAGCGGCGGTTCCGT	798
Qy	361	PheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro	380
Db	799	TGGGCAAAAGTC-----TCCATCGAAGCGGACACCAACGAA-----	834
Qy	381	AspLysLeuProValLysArgGluLeuGluGlnLeuThrValAsnMetGlyGlu	400
Db	835	-----GTCCCAACGCCGAACTGGGAAACTGCTGACCATGAAGCCCGCAAA	882
Qy	401	AlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe	420
Db	883	TGGTACGAACGCCGACGATGACCGCGTTTGGGT-----	918
Qy	421	AsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnVal	440
Db	919	-----GAGATTCAGAACCCGCGATGGGC	939
Qy	441	SerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu	460
Db	940	TCGGCAGGCTACGCATACAGC-----CAAAATCAGCGTA	972
Qy	461	GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro---	479
Db	973	CAGCGCTGCCGAAGCGCTGAAACCAACCGTCGATTTGCTCTGCATCAGCAACCGGC	1032
Qy	480	-----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu	492
Db	1033	CGAAATCTAGCTCAACGAATATACATCACCGCGCAACAAACCCCGCACGAA---	1089
Qy	493	AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp-----	508
Db	1090	---GTCGTCGCGCGTCAATTACGCCAAATGGAATCCGCACTTACGACACCTCCAAGCTG	1146
Qy	509	-----ArgValLeuAlaIleAsnHisAspGlyValAsnArgSerIle	523
Db	1147	CAACGTTCCAAAGAGCGCTCGAGCTTTTGGGCTACTTGCACAAATGTCCAG-----	1197
Qy	524	LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer	543
Db	1198	-----TTTGATGCTGCCCGCTTCCGCGGACGCCCGCACAA---	1233
Qy	544	GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla	563
Db	1234	-----GTCGATTTG-----AACATGAGTCTGACCGAAGCTTCCACC-----	1269
Qy	564	AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg	583
Db	1269	-----	1269
Qy	584	AspGlyGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr	601
Db	1270	---GTTCTCCTCGATTGAGCGCGGTGGTTCAAGATACCGGG-----TTGGTCATG	1320
Qy	602	LysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu	621
Db	1321	TCCGAGGCGTTTCCCAAGACAACCTGTTGCGGTACGGGCAAGTCGCGCCGACATCGCGCC	1380


```
Qy 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
Db 1381 TCCAGGACAAA-----ACCACGCTTAACGCGTCGCTGTGTTTACTGTACCGCTACTTC 1434
Qy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661
Db 1435 ACGCAGACGGGTCAGCGTGGGTAC-----CATGTTTACGGAANAAGCCTTCGACCCG 1488
Qy 662 PheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly 681
Db 1489 CGCAAGCATCGACCAAGCATCAAAATATAAACACCAACACCGCA-----GGC 1536
Qy 682 GlyTrpAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696
Db 1537 GCAGGCATCGCATGAGCGTCGCTTACCGCAATACGACCGCTGAATTTTCGTTGGTG 1596
Qy 697 LysThrGlnAlaProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713
Db 1597 GCAGAACACCTGACCGTCACACCTACACAAAGCGCCCAACACTATGCGACTTTATC 1656
Qy 714 Asn-----GlyLys-----ProSerGlnGlnAlaLeuAlaGlyValAlaValHisLys 730
Db 1657 AAGAAATACGCAAAACCGACGCGACAGCGAGCTTCAAAAGCTGCTGTACAAAGGT 1716
Qy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
Db 1717 ACCGTCGGTGGGGCGGCAACAAACCGACAGCGCGTTATGGCGGCGCGGTAC--- 1773
Qy 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
Db 1774 -----CTGACGGCGGTGAACGCCGAATC 1797
Qy 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
Db 1798 GCCCTGCTGGCAACAACTGCATATC-----TACTCGGCC 1833
Qy 785 AlaHisGlnMetThrGly----- 790
Db 1834 ACCCAACCAACCTCGTTCTTCCCTGTAGCAAAACCTTCACGCTGATGCTCGGGCG 1893
Qy 791 -----GlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArg 808
Db 1894 GAAGTCGGCATTCGGGCGCGCTAC-----GGCAGAACCAGAAATCCCTCTT 1944
Qy 809 LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
Db 1945 GAAACTTCTACGCGCGCGCGCTGGGTTTCGGTCGCGGATACGAAACGCGCACGCTCGGT 2004
Qy 829 ProfileSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly----- 845
Db 2005 CCG-----AAAGTCTATGAGATACGGCGAATAATCAGTACGGGGCAACAA 2055
Qy 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
Db 2056 AAAGCAACGCTCCGCGGAGCTGCTTCCCGATCCCGCGGCGAAGACGCGCGCAC 2115
Qy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
Db 2116 GTCGCGCTAGCGCTGTTTTCGCGAGCGAGCGGTGTTGGGACGCGCAAAACCTACGACGAC 2175
Qy 870 ----- 870
Db 2176 AACAGCAGTTCCGCGACCGCGCGGTTCAAAACATTTACGGCGCGCGCAATACCAT 2235
Qy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer 889
Db 2236 AAATCCACCTTTTACCAACGAATTCGGCTATTCCCGCGCGCGGTTCACCTGCTCG 2295
Qy 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
Db 2296 CTTTAGCCCGGATGAATTTACGTACGCTACCGCTAGGAAGAAAAACCGAAGACGAA 2355
Qy 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
```

Db 2356 ATCCAACGCTTCCCAATTCCAACTCGGCACGAGTTC 2391

RESULT 15

AAH42130
ID AAH42130 standard; DNA; 2394 BP.

XX AAH42130;

XX 17-SEP-2001 (first entry)

XX Nucleotide sequence of a Neisseria serogroup A protein.

XX Serogroup A protein; outer membrane protein; Neisserial infection;
KW vaccine; ss.

XX Neisseria meningitidis.

XX Key Location/Qualifiers

FT CDS 1..2394

FT /*tag= a

FT /product= "Neisseria serogroup A protein"

FT sig_peptide 1..63

FT /*tag= b

FT mat_peptide 64..2391

FT /*tag= c

PN WO200152885-A1.

XX 26-JUL-2001.

XX 17-JAN-2001; 2001WO-IB00166.

XX 17-JAN-2000; 2000GB-0001067.

XX 09-MAR-2000; 2000GB-0005699.

XX (CHIR-) CHIRON SPA.

XX Pizza M, Rappuoli R, Giuliani M;

XX WPI; 2001-451895/48.

XX P-PSDB; AAB84746.

XX Composition for treating or preventing infection to, detecting, or for
PT raising antibodies against Neisserial bacteria, comprises an N.
PT meningitidis serogroup B outer membrane preparation and an immunogenic
PT component .

XX Disclosure; Page 70-71; 83pp; English.

XX The present sequence encodes a Neisseria serogroup A protein. The
CC protein is used to produce the compositions of the invention. The
CC specification describes a composition, comprising a Neisseria
CC meningitidis serogroup B outer membrane preparation and an immunogenic
CC component. The immunogenic component is protein disclosed in WO99/57280,
CC WO99/36344, WO99/24578, WO99/66791, WO97/28273, WO96/29412, WO95/03413,
CC WO99/31132, WO99/58683, WO99/55873, and/or N. meningitidis protein PorA,
CC TbpA, TbpB, PilC, Opa, or Omp85. The composition is used for making
CC a medicament for treating or preventing infection due to Neisserial
CC bacteria; a diagnostic reagent for detecting the presence of Neisserial
CC bacteria or of antibodies raised against Neisserial bacteria; and/or
CC a reagent which can raise antibodies against Neisserial bacteria. It may
CC also be used as a vaccine.

XX SQ Sequence 2394 BP; 655 A; 758 C; 586 G; 395 T; 0 other;

Alignment Scores:

Pred. No.:	2,06e-11	Length:	2394
Score:	259.00	Matches:	191
Percent Similarity:	33.48%	Conservative:	121
Best Local Similarity:	20.49%	Mismatches:	342
Query Match:	5.48%	Indels:	278
DB:	22	Gaps:	40

US-09-914-168-2 (1-919) x AAH42130 (1-2394)

Qy 116 ThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
Db 46 TCGGCTTTGGCAGCTTGGCGACTTCCACCTCAACCAAGGTCGCGTCAAGGTCGAGCGT 105
Qy 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149
Db 106 ACCGAGCGGAGTACCGTATTCAACTACCTGCCGCTCAAGTCGGGACACCTACAAAC--- 162
Qy 150 GluValValProProThrLeuGluProGluLysProGlyLeuIleAspArgLeuTyr 169
Db 163 -----GACACACGCGAGTGCATCATCAAAAGCCTGTAC 198
Qy 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
Db 199 GCCACCGGTTCTTGGACGAGTACCGCTCGAAGCTCGGACGGCGGACGCTCTCGTGACC 258
Qy 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
Db 259 GTTATCGAAGCCGCCACCATCGCTCGCTCAACATCACCGCGCAAAATGCTGCAAAAC 318
Qy 207 GluProTyrAlaAsnIleLysAlaLeuGluAsp-----IleThrGlnGluSerAla 224
Db 319 GACGCC-----ATTAAAGAAAACCTCGAATCGTTCGGGCTGGCGCAGTCGCAATAC 369
Qy 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
Db 370 TTTAATCAGGCGACACTCAATCAGCAGTCGCGCGCTGGAAGAGAATACCTCGGGGCGC 429
Qy 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
Db 430 GCGAAACTCAATATCCAAATCAGCCCAAGTAACCAAACTCGCCCGCAACCGCTCGAC 489
Qy 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
Db 490 ATCGACATCAGCATGACGAGGCGAAATCCGCCAAATCAACCGACATCGAATTTGAA--- 546
Qy 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluAlaAspAsp 289
Db 547 GCGAACCAAGTCTATTCGACCGGCAAACTGATGCGG----- 582
Qy 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
Db 583 -----CAGATGTCGCTGACCGAAGGCGCATTTGGACATGG 618
Qy 307 -----PheHisGlyLysTyrGluThrLysLysAsnLeuIleGlu 320
Db 619 CTGACACGAAGCAACCAATTCACGAGCAGAGAATTTGCCCAAGACATGGAAAGTAACC 678
Qy 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrLeuAspArgSerValAspVal 340
Db 679 GACTTCTACGAGAACAGCGGTACTTCGATTCCCGCATCTTCGATACCGACATCCCAACC 738
Qy 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArg 360
Db 739 AACGAAGCAAAACCAAGCAGACCATCAAAATCACCGTCCACGAAGGCGGAGTTTCGCT 798
Qy 361 PheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
Db 799 TGGGCAAAAGTC-----TCCATCGAAGCGCGACACCAAGAA----- 834
Qy 381 AspLysLeuProValLysArgGluLeuGluGlnLeuLeuThrValAsnMetGlyGlu 400
Db 835 -----GTCCCAAGCCCACTGGAAAACCTGCTGACCATGAAGCCCGGCAAA 882
Qy 401 AlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
Db 883 TGGTACGAAGCCGACGAGTACCGCGCTTTTGGGT----- 918
Qy 421 AsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnVal 440
Db -----

Db 919 -----GAGATTTCAGAACCGCATGGC 939
Qy 441 SerPheGlnGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu 460
Db 940 TCGGACAGCTACGCATACAGC-----GAAATCAGCGTA 972
Qy 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspLysSerPro--- 479
Db 973 CAGCGCTGCCAAACCGCGAAACCAACCGCTGATTCGTCTGCATCAACACCGGCGC 1032
Qy 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492
Db 1033 CGAAATTCACGTCACCAACCAATCCACATCACCGCAACCAACCAACCGCGACGAA--- 1089
Qy 493 AsnLeuValAlaLysAlaArgHisLeuTyrAspMetProAspAsp----- 508
Db 1090 ---GTCGTGCGCGGCAATTCGCCCAATGGAATCGCGCCTTACGACACCTCCAAGCTG 1146
Qy 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle 523
Db 1147 CAACGCTCCAAAGAGCGCTCGAGCTTTTGGGCTACTTCGACACGCTACAG----- 1197
Qy 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543
Db 1198 -----TTTGATGCGCTCCCGCTTGGCGGCACACCCGACAAA--- 1233
Qy 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
Db 1234 -----GTCGATTTG-----AACATGAGCCTGACCGAAGCTTCACC----- 1269
Qy 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583
Db 1269 ----- 1269
Qy 584 AspGlyGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr 601
Db 1270 ---GGCTCGCTCGACTTGAGCGGCGTGGGTACAGGATACCGGC-----CTGGTCATG 1320
Qy 602 LysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 1321 TCCGAGGCGTTTCCCAAGACAACTGTTCGTACGGCAAGTCGCGCCCTCGCGCC 1380
Qy 622 SerGluAspLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
Db 1381 TCACGAAGCAAA-----ACCACGCTCAACGCTCGCTCGCTTTACCGACCGTACTTC 1434
Qy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661
Db 1435 ACGGACAGCGGCTCAGCTGGGCTAC-----GATGTTTACGGAAGGCTTCGACCCG 1488
Qy 662 PheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly 681
Db 1489 CCGAAGCATCCACGACATCAACAATAATAAACACCAACCGCGCA-----GGC 1536
Qy 682 GlyTyrAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696
Db 1537 GCAGGCATCCGATGAGCGTGTGTACCGAATACGACCGCGTGAATTCGTTGGTGTG 1596
Qy 697 LysThrGlnAlaProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713
Db 1597 GCAGAACACCTGACCGTCAACACCTACAAACAAAGCGCCCAACACATATGCCACTTATC 1656
Qy 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730
Db 1657 AGAATACGGCAAAACCGCAGCGCAGCGAGCTTCAAAGGCTGGCTGTACAAGGT 1716
Qy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgIleTyrArg 744
Db 1717 ACCGTGCGGTGGGGCGCAACAAACCGACAGCGCTTATGGCGCGACGCGGTAC--- 1773
Qy 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
Db 1774 -----CTGACGGCGGTGAACCGCGAATC 1797

```
QY 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
DB 1798 GCCCTGCCGCGCAGCAAACTGCAATAC-----TACTCGGCC 1833
QY 785 AlaHisGlnMetThrGly----- 790
DB 1834 ACCCACACCAACCACTGGTCTCCCTTAAAGCAAACTTCACGCTGATGCTCGGCGGC 1893
QY 791 -----GlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArg 808
DB 1894 GAAGTCGGCATTCGCGCGCGCTAC-----GGCAGAACCAAGAAATCCCTTCTTT 1944
QY 809 LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
DB 1945 GAAAACTTTCACGCGCGCGCTGGTTCGGTGGCGGATACGAAAGCGGCACGCTCGGT 2004
QY 829 ProIleSerAspGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly----- 845
DB 2005 CCG-----AAAGTGTATGACGAATACGGCGAAATAATCAGCTACGGCGGCAACAA 2055
QY 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
DB 2056 AAGCCCAACGTCTCCGCGAGCTCTCTCCGATGCCGCGGCGAAAGCGCGCAC 2115
QY 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
DB 2116 GTCGCGCTGACGCTGTTCGCGAGCGAGCGAGCGTGTGGAGCGCAAAACCTAGACGAC 2175
QY 870 ----- 870
DB 2176 AACAGCAGTTCGCGCGCGCGAGGTTCAAAACATTTACGGCGCGGCAATACCAT 2235
QY 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTtrPAlaSer 889
DB 2236 AAATCCACCTTTACCAACGAATTCGCTATTCCGCGCGCGCGGTTACTGCTCTCG 2295
QY 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
DB 2296 CTTTAGCGCGGAGGAATTCAGCTAGCGCTACCGCTGAGAAANAAACCGAAGACGAA 2355
QY 908 ProIleLysLeuHisPheIleGlyThrProPhe 919
DB 2356 ATCCAACGCTTCCAATTCCAACTCGGCACGACGCTC 2391
RESULT 16
AA81478
ID AAA81478 standard; DNA; 52253 BP.
AC AAA81478;
XX
DB 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_26 SEQ ID NO:26.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN W0200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-US233573.
XX
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX
PA (CHIR ) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
```

```
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
DR WPI; 2000-318079/27.
XX
PT Isolated nucleotide sequences of Neisseria meningitidis which can be
PT used in the diagnosis and treatment of N. meningitidis infection and
PT other Neisserial infections, for example, N.gonorrhoea -
XX
XX Claim 7: Page 532-547; 1760pp; English.
XX
CC The present invention describes methods of obtaining immunogenic
CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
CC represent specifically claimed Neisseria meningitidis genomic DNA
CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to
CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
CC sequences, which are all used in the exemplification of the present
CC invention. The nucleic acid sequences, protein sequences, and antibodies
CC against them, can be used in the manufacture of a composition. The
CC composition can be used as a medicament (or in the manufacture of a
CC medicament) for treating, preventing or diagnosing infection due to
CC Neisserial bacteria. For example, some of the identified proteins could
CC be components of vaccines against Meningococcus B; against all serotypes;
CC and/or against all pathogenic Neisseriae. Identification of sequences
CC from the bacterium will also facilitate production of biological probes,
CC particularly organism-specific probes. Attempts to make efficacious
CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
CC Multivalent vaccines have also been tried but none have successfully
CC overcome antigenic variability. The provision of further, complete
CC sequences may provide an opportunity to identify secreted or surface
CC exposed proteins that may be presumed targets for the immune system and
CC which are not antigenically variable or at least more conserved than
CC other more variable regions.
XX
```

Sequence 52253 BP; 12877 A; 15510 C; 13103 G; 10761 T; 2 other;

Alignment Scores:

Pred. No.:	1.17e-09	Length:	52253
Score:	259.00	Matches:	191
Percent Similarity:	33.48%	Conservative:	121
Best Local Similarity:	20.49%	Mismatches:	342
Query Match:	5.48%	Indels:	278
DB:	21	Gaps:	40

US-09-914-168-2 (1-919) x AAA81478 (1-52253)

```
QY 116 ThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
DB 47343 TCGCCTTTGGCAGCTTGCAGACTTCACCATCCAGACATCCCGCTCGAAGGCTTGACGGT 47402
QY 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149
DB 47403 ACCGAGCGGAGTACCGGTATTCAACTACTGCGCGTCAAAAGTCGGCGACACCTCAAC--- 47459
QY 150 GluValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
DB 47460 -----GACACACACGCGAGTGCATCATCAAAAGCCTGTAC 47495
QY 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
DB 47496 GCCACCGGTTCCTTTGTGACGAGTACCGGTCCGAAACTCCGCGGCGAGCTCTCTCTGAC 47555
QY 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
DB 47556 GTTATCGAACGCCCCACCATCGGCTCGCTCAACATCACCGCGCGCAAAANCTCTGCAAAAC 47615
QY 207 GluProTyrAlaAsnIleLysAlaAlaLeuGluAsp-----IleThrGlnGluSerAla 224
DB 47616 GACGCC-----ATTAAGAAAAAGCTCGAATCGTTCGGGCTGGCGCAGCTCGCAATAC 47666
```

```
QY 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
Db 47667 TTTAATCAGCGCACACTCAATCAGGCAGTCGCGCGCTGAAAAGAAATACCTCGCGCGC 47726
QY 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrArgPhe 249
Db 47727 GGCAAACTCAATATCCAAATCAGCCCAAGTAACCAACTCGCCGCAACCGCTCGAC 47786
QY 250 IleAspLeuSerIleIleArgAsnSerIleGlyValAlaAspValIleIleHisAspLeu 269
Db 47787 ATCGACATCAGCATTCAGCAGGCGCAATCCGCCAAATCACCACATCGAATTGAA--- 47843
QY 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289
Db 47844 GCAACCAAGTCTATTCGCCAGCGCAAACTGATCGCG----- 47879
QY 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
Db 47880 -----CNAATGTCCTGACCGCAAGCGCGCAATTTGGACATGG 47915
QY 307 -----PheHisGlyLysTyrGluThrLysLysAsnLeuIleGlu 320
Db 47916 CTGACACGAAGCAACCAATTCACGAGCAGAAATTTGCCCAAGATATGAAAAGTAACC 47975
QY 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPheAspArgSerValAspVal 340
Db 47976 GACTTCTACCAAAATAACGGCTACTTCGATTTCCGTATCTCGATACCGACATCCAAACC 48035
QY 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArg 360
Db 48036 ACGAAGCAAAACCAAGCAGACCATCAAAATCACCGTCCAGAGCGCGAGCTTTCCTG 48095
QY 361 PheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
Db 48096 TGGGGCAAGTC-----TCCATCGAAGCGCACACCAACGAA----- 48131
QY 381 AspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGlu 400
Db 48132 -----GTCCCCAAACCGCAACTGGAAAACTGCTGACCATGAAGCCCGCAAA 48179
QY 401 AlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
Db 48180 TGGTACGAACGCCAGCAGATGACCGCGTTTGGGT----- 48215
QY 421 AsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnVal 440
Db 48216 -----GAGATTCAAGAACCGCATGGGC 48236
QY 441 SerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu 460
Db 48237 TCGGAGCGTACGCATACAGC-----GAAATCAGCGTA 48269
QY 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro--- 479
Db 48270 CAGCGCTGCCAGCGCTGAACCAACACCGTCGATTTCGTCTGCATCGAACCGGC 48329
QY 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492
Db 48330 CGGAAATCTAGCTCAACGAAATACATCACCGGCAACAAACAAACCGCGCACGAA--- 48386
QY 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp----- 508
Db 48387 ---GTCGTCGCGGTGAATTACCGCAAAATGGAAATCCGCACTTACGACCTCCCAAGCTG 48443
QY 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAlaAsnArgSerIle 523
Db 48444 CAACGTTCCAAAGCGGCTCAGCTTTTGGGCTACTTCGACAAATGTCAG----- 48494
QY 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543
Db 48495 -----TTTCATGCTGCTCCCGCTTGGCGGCGCGCGCACGCAAA--- 48530
QY 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
Db 48531 -----GTCGATTG-----AACATGAGTCTGACCGAAGCTTCCAC----- 48566
QY 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583
Db 48566 ----- 48566
QY 584 AspGlyGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr 601
Db 48567 ---GGTTCCCTGGATTGAGCGCGGTGGTTCAAGATACCGG-----TTGGTTCATG 48617
QY 602 LysPheGluHisAsnLeuLeuAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 48618 TCCGCAAGGCGTTTCCCAAGACACACCTGTTCCGTACGGCAAGTCGCGCGCACTCGCGCC 48677
QY 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
Db 48678 TCCAGAGCAAA-----ACCAGCTTAACGGCTCCGCTGCTGTTTACTGACCGCTACTTC 48731
QY 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPheGlyHisSerThrAsnGly 661
Db 48732 ACGGCAGACGGGTGACGCTGGGCTAC-----GATGTTTACGGAAGAGCTTTCGACCCG 48785
QY 662 PheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly 681
Db 48786 CGCAAGACATCGACACGATCAACAATATAAAACCAACCAACGCA-----GGC 48833
QY 682 GlyTyrAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696
Db 48834 GCAGGCATCGCATGAGCGTGCCTGTTCCGAATACAGCCGCTGAATTCGCTTGGTGGTG 48893
QY 697 LysThrGlnAlaProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713
Db 48894 GCAGAACACCTGACCGCTCAACACCTACACAAAGCGCCCAACACATATGCGCATTTATC 48953
QY 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730
Db 48954 AAGAAATACGGCAAAACCGCACGACGCGAGCTTCAAAGGCTGGGTGTACAAAGGT 49013
QY 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
Db 49014 ACCGTGCGTGGGGCGCAACAAACCGACAGCGCGTTATGGCCGCGCGCTAC--- 49070
QY 745 GlnArgTyrSerLeuGluValGlySerGlyLeuValSerAspAlaAsnMetAlaIle 764
Db 49071 -----CTGACGGCGTGAACGCGCAAAATC 49094
QY 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
Db 49095 GCGCTGCGTGGCAGCAAACTGCAATAC-----TACTCCGCC 49130
QY 785 AlaHisGlnMetThrGly----- 790
Db 49131 ACCCAACCAACCACTGGTTCTTCCCTCGACCAAAACCTTACGCTGATGCTCGCGGC 49190
QY 791 -----GlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArg 808
Db 49191 GAAGTCGGCATTCGGCGCGCTAC-----GGCAGAACCAAGAAATCCCTCTCTTT 49241
QY 809 LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisaspSerLeuSer 828
Db 49242 GAAACTTCTACGGCGCGCGCTGGGTTCGTCGCGGATACGAAAGCGCGCACGCTCGGT 49301
QY 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly----- 845
Db 49302 CCG-----AAAGTCTATGACGAAATACGGGAAAAAATACGTACGCGCGCAACAA 49352
QY 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
Db 49353 AAAGCAACGCTCTCCGCGCGAGCTGCTCTCCCGATCCCGCGCGGCGGCGCGCAC 49412
QY 857 LeuArgLeuAlaPheGlyAspIleGlyAsnAlaTyrAsp----- 870
Db 870 ----- 870
```

Db 49413 GTCCGCTGACCTGTTTGGCCGACGACGAGCGTGTGGGAGCGGCAAAACCTACGACGAC 49472
Qy 870 ----- 870
Db 49473 AACAGCAGTTCGCGACGCGCGGAGGTTTCAAAACATTTTACGCGCGGCAATACCCAT 49532
Qy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer 889
Db 49533 AAATCCACCTTTTACCAAGAAATTCGCTATTCCGCGCGCGGCTTACCTGCTCTCG 49592
Qy 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
Db 49593 CCTTAGCCCGATGAATTCAGTACGCTACCGCTGAAGAAAAACCGAGACGAA 49652
Qy 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
Db 49653 ATCCAAGCTTCCAATTCCTCAACTCGGACGAGGTTTC 49688
RESULT 17
AAF21544/c
ID AAF21544 standard; DNA; 349980 BP.
XX
AC AAF21544;
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:1.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX
OS Neisseria meningitidis.
XX
PN WO200066791-A1.
XX
XX 09-NOV-2000.
XX
PD 08-MAR-2000; 2000WO-US05928.
PF
XX 30-APR-1999; 99US-0132068.
PR 08-OCT-1999; 99WO-US23573.
PR 28-FEB-2000; 2000GB-0004695.
XX
XX (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
XX Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Massignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;
XX
DR WPI; 2000-647603/62.
XX
XX Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections -
PT
XX
PS Claim 7; Appendix A; 692pp; English.
XX
XX The present invention describes the full length genome of
CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
CC to AAF21613 represent fragments of the NMB genomic sequence, as the
CC sequence was too long to go in a record on its own it was split into 8
CC sequences which overlap each other at the beginning and end of each
CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
CC AAF21606 represent PCR primers which are used in the exemplification of
CC the present invention. The NMB genome and fragments from it have
CC antibacterial activity, and can be used in vaccines and gene therapy.
CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
CC proteins can be used in compositions for treating or preventing infection
CC due to Neisserial bacteria or as a diagnostic reagent for detecting the

CC presence of Neisserial bacteria or of antibodies raised to Neisserial
CC bacteria. Computers, computer memory, computer storage medium or computer
CC databases can be used in a search to identify open reading frames (ORFs)
CC or coding sequences within the NMB genome. The DNA sequences provide
CC further opportunities to find antigenic or immunogenic proteins which are
CC more effective in vaccines than the outer membrane proteins currently
CC used.
XX
SQ Sequence 349980 BP; 83241 A; 85091 C; 95206 G; 86442 T; 0 other;
Alignment Scores:
Pred. No.: 1.41e-08 Length: 349980
Score: 259.00 Matches: 191
Percent Similarity: 33.48% Conservative: 121
Best Local Similarity: 20.49% Mismatches: 342
Query Match: 5.48% Indels: 278
DB: 21 Gaps: 40
US-09-914-168-2 (1-919) x AAF21544 (1-349980)
Qy 116 ThrProLeuSerLeuGluGluLeuPheAlaGlnGlnSerThrGluMetGlyIleAsnPro 135
Db 180761 TCGCCTTTGGCACTTGGCGACTTCACCATCCAAGACATCCGCGCTCGAAGGCTTCGAGCGT 180702
Qy 136 AsnAsp-----TyrIleProGluTyrGlnGlyGluGlnProAsnSer 149
Db 180701 ACCGACGCGAGTACGCTATTCAACTACCTGCCCGTCAAGTCGGCGACACCTACAAAC--- 180645
Qy 150 GluValValProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
Db 180644 -----GACACACAGCGCAGTGCATCATCANAAGCCCTGTAC 180609
Qy 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
Db 180608 GCCACCGGTTTCTTTGACGAGCTACGCGTCAAGAACTGCGGAGCGGCGAGCTCTGCTGACC 180349
Qy 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
Db 180548 GTTATCGAAGCGCCGCCACCATCGCTCGCTCAACATCACCGCGCAAAATATGCTGCAAAAC 180489
Qy 207 GluProTyrAlaAsnIleLysAlaLeuGluAsp-----IleThrGlnGluSerAla 224
Db 180488 GAGGCC-----ATTAAGAAAAACCTCGAATCTTTCGGGCTGGCGCAGTGCGAATAC 180438
Qy 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
Db 180437 TTTAATACGCGCACACTCAATCAGCGAGTCCGCGCGCTGAAAGAAAGAAATACCTCGGCGCC 180378
Qy 237 -----ThrAlaLeuValAlaAlaArgAlaValAlcIlyTyrTyrAsp 249
Db 180377 GGCAAACTCAATATCCAATCAGCCCAAGTAACCAACTCGCCCGCAACCGCGCTCGAC 180318
Qy 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
Db 180317 ATCGACATCAGCATTTGACGAGGGCAAAATCCGCCAAATACCCGACATCGAATTTGAA--- 180261
Qy 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289
Db 180260 GGCAACCAAGTCTATTTCGCGCGCAAACTGATGCGG----- 180225
Qy 290 LysAlaPheThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
Db 180224 -----CAAATGTCCCTGACCGAGGCGGCGATTTTGGACATCGG 180189
Qy 307 -----PheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGlu 320
Db 180188 CTGACACGAACCAACCAATTCACCGCAGAGAAATTTGCCCAAGATATGAAAAAAGTAACC 180129
Qy 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPheLeuAspArgSerValAspVal 340
Db 180128 GACTTCTACCAAAAATACGGGCTACTTCGATTTCCGCTCGATACCGCAGATTCACCAACC 180069
Qy 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArg 360

Db 180068 AACGAAGCAAAACCAAGCAGACATCAAAATCACCGTCCACGAAGCGGAGCTTCCGT 180009
Qy 361 PheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
Db 180008 TGGGCAAGTC-----TCCATCGAGCGGCACCAACGAA----- 179973
Qy 381 AspLysLeuProValLysArgGluLeuLeuGluLeuLeuValAsnMetGlyGlu 400
Db 179972 -----GTCCCAAAAGCCGAACCTGGAAACCTGCTGACCATGAAGCCCGCAA 179925
Qy 401 AlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
Db 179924 TGGTACGAACGCAGCATGACCGCGTGTGGGT----- 179889
Qy 421 AsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnVal 440
Db 179888 -----GAGATTCAGAACCCGATGGGC 179868
Qy 441 SerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu 460
Db 179867 TCGGAGGCTACGCATACAGC-----GAAATCAGCGTA 179835
Qy 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro--- 479
Db 179834 CAGCGCTGCCGACGCTGAACCAACCGTTCGATTTCGTCTCCATCGAACCCGGGC 179775
Qy 480 -----IleGluPheSerAlaSerAlaSerLeuIleGlnAspLysLeu 492
Db 179774 CGGAAATCTAGTCAACGAATATACATCACCGCGCAACAAACCCGCGACGAA--- 179718
Qy 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp----- 508
Db 179717 ---GTCGTCGCGCGTGAATTCGCAAAATGGAAATCCGACCTTACGACCTCCCAAGCTG 179661
Qy 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle 523
Db 179660 CAACGTTCCAAAGCGGCTCGAGCTTTGGGCTACTTCGACAATGTCAG----- 179610
Qy 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543
Db 179609 -----TTTGATGCTGTCGCCGCTGCGCGCACGCCGACAAA--- 179574
Qy 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
Db 179573 -----GTCGATTTG-----AACATGAGTCTGACCAACGTTCCAC--- 179538
Qy 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583
Db 179538 ----- 179538
Qy 584 AspGlyGlnIleGlyLeu-----GlyTrpGlySerAspThrGlyThrArgLeuValThr 601
Db 179537 ---GTTCCCTCGATTGAGCGCGGTTGGGTTCAGATACCGGG-----TTGTCATG 179487
Qy 602 LysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 179486 TCCGACGCGGTTTCCCAAGACAACCTGTTCCGAGCGCAAGTCGCGCGCATCTGCGCGCC 179427
Qy 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
Db 179426 TCCAGAGCAMA-----ACCAGCTTAACGGCTCGCTGCTGCTTACTGACCCGACTTTC 179373
Qy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661
Db 179372 ACGGCAGACGGGTCAGCTGGGCTAC-----GATGTTTACGGAAGACCTTCGACCCG 179319
Qy 662 PheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly 681
Db 179318 CGCAAGCATTCGACGAGCATCAACAAATATAAAACCAACCAACGCA-----GGC 179271
Qy 682 GlyTyrAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696

Db 179270 GCAGGCATCCGATGAGCGTGTGCTTACCGAATACGACCGCTGAATTTCCGTTGGTG 179211
Qy 697 LysThrGlnAlaProProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713
Db 179210 GCAGAACACCTGACCGTCAACACCTACAAACAAAGCCGCCAAACACTATGCCGACTTATC 179151
Qy 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730
Db 179150 AAGAAATACGCCAAACCCGACGACGACGACGCTTCAAAGGCTGGCTGTACAAAGGT 179091
Qy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
Db 179090 ACCCTGCGCTGGGCGCGCAACAAACCGACAGCGGTATGCGCGACGCGGCTAC--- 179034
Qy 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
Db 179033 -----CTGACGGCGGTGAACCCCGAATC 179010
Qy 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
Db 179009 GCCCTGCTGCGACGAAACTGCAATAC-----TACTCCGCG 178974
Qy 785 AlaHisGlnMetThrGly----- 790
Db 178973 ACCCAACCAACCTGGTCTTCCCTCGACAAACCTTCCGCTGATGCTCGCGGC 178914
Qy 791 -----GlyIleGlnAlaGlyTyrIleTrpSerAspAsnHisValProTyrArg 808
Db 178913 GAAGTCGGCATTCGCGCGGCTAC-----GGCAGAACCAAGAAATCCCTTCTTT 178863
Qy 809 LeuArgPheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
Db 178862 GAAAACTTCTACGCGCGCGCTGGGTTCGTCGCGGATACGAAAGCGCAGCTCGGT 178803
Qy 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly----- 845
Db 178802 CCG-----AAAGTCTATGACGAATACGGCGGAAATAATCAGCTACGCGCGCAACAA 178752
Qy 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
Db 178751 AAAGCAAGCTCTCCGCGAGCTGCTTCCCGATGCCGCGCGCAAGACGCGCGCAC 178692
Qy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
Db 178691 GTCCGCTGAGCTGTTGCGCAGCAGCAGCGCTGTGGGAGCGGCAAAACCTACGACGAC 178632
Qy 870 ----- 870
Db 178631 AACACAGTTCCGCGACGCGCGGAGGGTTCAAAACATTTACGGCGCGCAATACCCAT 178572
Qy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer 889
Db 178571 AAATCCACCTTTACCAACGAATTCGCTATTCCGCGCGCGCGGCTTACCTGGCTCTCG 178512
Qy 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
Db 178511 CCTTAGCGCCGATGAAATTCAGCTACGCTACCGCTACCGCTGAAGAAAAACCGGAAGACGAA 178452
Qy 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
Db 178451 ATCCAACGCTTCCATTTCCAATTCGACGCGACGAGCTTC 178416
RESULT 18
ID AAA15156 standard; DNA; 2394 BP.
XX
AC AAA15156;
XX
DT 21-AUG-2000 (first entry)
XX
DE DNA encoding outer membrane protein (omp) 85.
KW Outer membrane protein; omp85; gonococcal infection; symptomatic disease;

Db 1147 CAACGCTCCAAAGACGCGTTCGAGCTTTTGGCTACTTCGACACGTACAG----- 1197
 Qy 524 LeuGlyArgIleSerAspAlaValIleSerAlaValAlaArgAlaIleLeuProAspGluSer 543
 Db 1198 -----TTTGATGCGCTCCGCTTGGCGGACACCGACAAA--- 1233
 Qy 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
 Db 1234 -----GTGATTTG-----AACATGAGCTGACCGAACGTTCCACC----- 1269
 Qy 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583
 Db 1269 ----- 1269
 Qy 584 AspGlyGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr 601
 Db 1270 ---GGCTCGCTCGACTTGAGCGCGGCTGGGTACAGGATACCGC-----CTGTCATG 1320
 Qy 602 LysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
 Db 1321 TCCGACGGCGTTTCCCAAGACAACTGTTCGTACGGGCAAGTCGGCGCCTGCGGCGC 1380
 Qy 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
 Db 1381 TCACGAAGCAAA-----ACCACGCTCAACGCGCTCGCTGCTTTACCGACCCGTACTTC 1434
 Qy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661
 Db 1435 ACGCAGACGGGTGAGCTGGCTAC-----GATGTTTACGGAAGCCTTCGACCGC 1488
 Qy 662 PheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGly 681
 Db 1489 CGCAAAGCATCGACGACGATCAACAATATAAAACCAACACCGGCA-----GGC 1536
 Qy 682 GlyTrpAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696
 Db 1537 CGAGCATCCGATGAGCGTCCCTGTTACCGAATACGACCGCGTGAATTCGGTTTGGTG 1596
 Qy 697 LysThrGlnAlaProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713
 Db 1597 GCAGAACACCTGACCGTCAACACCTCAACAAAGCGGCCAACACTATGCGGACTTTATC 1656
 Qy 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730
 Db 1657 AAGAAATACGCAAAACCGACGCGACGCGACGCGAGCTTCAAGGCTGCTGTACAAAGGT 1716
 Qy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
 Db 1717 ACCGTGCGCTGGGCGGCAACAAACCGACGCGGTATTGCGCGCGCGGTAC--- 1773
 Qy 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
 Db 1774 -----CTGACGGCGTCAACGCCGGAATC 1797
 Qy 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
 Db 1798 GCCCTGCCGCGCAAACTGCAATAC-----TACTCCGCC 1833
 Qy 785 AlaHisGlnMetThrGly----- 790
 Db 1834 ACCCACACCAAACTGGTTCTTCCCTTAAGCAAAACCTTCACGCTGATGCTCGGCGGC 1893
 Qy 791 -----GlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArg 808
 Db 1894 GAAGTCGCAATGCGGCGGCTAC-----GCACGAACCAAGAAATCCCTTCTTT 1944
 Qy 809 LeuArgPheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLysSer 828
 Db 1945 GAAAACTTCTACGGCGCGGCGCTGGGTTCGGTTCGGGATACGAAAGCGGACGCTCGGT 2004
 Qy 829 ProfileSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly----- 845
 Db 2005 CCG-----AAAGTGTATGACGAATACGGGAAAAAATCAGCTACGCGGCGCAACAA 2055

Qy 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
 Db 2056 AAGCCCAACGCTCCGCGGAGCTGCTCTCCGATGCTCGGCGGAAGACGCGCGCAC 2115
 Qy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
 Db 2116 GTCCGCTGAGCCTGTTTTCGCGACGAGCAGCGCTGTGGGACGGAACCTTACGACGAC 2175
 Qy 870 ----- 870
 Db 2176 AACAGCAGTTCCGCGACGCGGCGAGGTTCAAAACATTTACGGCGCGGCANTACCCAT 2235
 Qy 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer 889
 Db 2236 AAATCCACCTTTACCAACGAATTCGCTATTCCGCGCGCGCGGCTTACCTGGCTCTCG 2295
 Qy 890 ProValGlyGlnValArgValaspValAlaThrGlyValLysGlu-----GluGlyAsn 907
 Db 2296 CTTTAGCCCGCGATGAATTCAGGTACGCTACCGCTGAAGAAAAAACCGGAAGACGAA 2355
 Qy 908 ProfileLysLeuHisPhePheIleGlyThrProPhe 919
 Db 2356 ATCCAACGCTTCCATTCCAACTCGGACGACGCTTC 2391

RESULT 19

AAI97964
 ID AAI97964 standard; DNA; 6617 BP.
 XX
 AC AAI97964;
 DT 20-NOV-2001 (first entry)
 XX
 DE Lawsonia intracellularis coding sequence SEQ ID NO: 1.
 XX
 KW HtrA; PonA; HycP; YefW; ABC1; Omp100; Lawsonia intracellularis infection;
 KW vaccine; ds.
 XX
 OS Lawsonia intracellularis.
 XX
 PN JP2001169787-A.
 XX
 PD 26-JUN-2001.
 XX
 PF 20-OCT-2000; 2000JP-0320736.
 XX
 PR 22-OCT-1999; 99US-0160922.
 XX
 PA (PFIZ) PFIZER PROD INC.
 XX
 DR WPT; 2001-592540/67.
 XX
 PT Lawsonia intracellularis polynucleotide and encoded protein, used to
 PT prevent Lawsonia intracellularis infection -
 XX
 PS Claim 3; Page 37-39; 67pp; Japanese.
 XX
 CC The present invention provides isolated polynucleotides encoding HtrA,
 CC PonA, HycP, LysS, YefW, ABC1 or Omp100 protein of Lawsonia
 CC intracellularis. The sequences can be used in vaccines for the prevention
 CC of Lawsonia intracellularis infection. The present sequence is a coding
 CC sequence of the invention.
 XX
 SQ Sequence 6617 BP; 2216 A; 1153 C; 1142 G; 2106 T; 0 other;

Alignment Scores:

Pred. No.: 3.65e-09 Length: 6617
 Score: 238.50 Matches: 198
 Percent Similarity: 34.72% Conservative: 137
 Best Local Similarity: 20.52% Mismatches: 391
 Query Match: 5.05% Indels: 239
 Gaps: 38

US-09-914-168-2 (1-919) x AAI97964 (1-6617)

Qy	50	AsnGlnAlaLysAlaGlyAsnProProValLeuLeuLeuThrProGluGlnIleGlnAlaArg	69
Db	3764	ANTGCTGCTTCAAAGACGACTCTCTATTTGGTGTCTCCCAATTTCAATTAATGGTGCA	3823
Qy	70	LeuAsnAlaAlaGlyLeuAsnAlaLys---ProGlnSerGlnAlaLeuAspValValAsn	88
Db	3824	TCAANTGATGAAGAGTTACAAACAGAACTACCACATGCTTCTTGCACTGCATTAAGAAT	3883
Qy	89	PheAspAspGlnSerProIleSerArgIleGlyGluGlnSerProProLeuGlyLeuAsp	108
Db	3884	-----AAGGGA	3889
Qy	109	MetSerValIleGluGluThrThrProLeuSerLeuGluGluLeuPheAlaGlnGluSer	128
Db	3890	TTTGGTGTTCCTCCTAATAATTCGATTAATATTT-----CTATATAACAATAATC	3943
Qy	129	ThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnProAsn	148
Db	3944	TCCCAACTTAATTTCTACTCAAAAAGGTA-----GCTCAACACTCCAT	3991
Qy	149	SerGluValValValProThrLeuGluProGluLysProGlyLeuIleLysArgLeu	168
Db	3992	GCTGACTATCTAGTA-----	4006
Qy	169	TyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyr	188
Db	4007	TACGGCAGTTTCANTCAACAGGTGAAT-----TTTAGT	4042
Qy	189	GlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSer-----HisGlnLys	205
Db	4043	ATTGATAGTAGGCTTATTGATAGTACAGGTGTAGCATCTGCACGTCCCATATATAGAA	4102
Qy	206	ThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAlaMet	225
Db	4103	AAACCAAAATTTAATGAGCTAAATATTGCTGTGAACAGAACTTGCTGACGTATAAGT---	4159
Qy	226	AspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaVal	245
Db	4160	-----AATGCCTTATAAGAAA---AACACTATTGCTGATGATACGTATTTCATGGGCTT	4210
Qy	246	GlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIle	265
Db	4211	AAAGTTCTTGATCCTGATGATTCCTTACACGACTCACTATTAATAAAGGGAGATCATACT	4270
Qy	266	IleHis-----AspLeuGlyGluProValTyr	274
Db	4271	GATCATCCCAAAATTAATGCAGAAATCAAAAATATGGAATTAGA-----TATTTT	4324
Qy	275	IleAspTyrArgAla---ValGluValArgGlyGluGlyAlaAspAspLysAlaPheThr	293
Db	4325	AGTGATCTCTGCAAGTATTGAAGAAAGCGGGGAAGA-----CGATTACTTGTA	4375
Qy	294	ThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisGlyLysTyrGlu	313
Db	4376	TTTACTGTACAGAAGAGCTAAAATACAGATGTGTGTTCAGCGCTCAAAAGCTGTA	4435
Qy	314	ThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyr	333
Db	4436	AGTATCGATAACATTTCTGCTGCAATGATTTCTAAAAAAGGA-----	4477
Qy	334	LeuAspArgSerVal-----AspValIleLeuProAspAsnThrAlaAspValSerLeu	351
Db	4478	-----TCAGTTATTAGTAGACATATTGTCCTCCAGATATTTCAAAAATATTCACGAC	4528
Qy	352	IleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIleAspPro	371
Db	4529	CTCTATAGAAAGAGGCTACTATCTCGCTGAA-----GTTAATTATGAATAAAGAG	4582
Qy	372	LysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGlu	391
Db	4583	AAAGAAATACTTCTTCTGCA-----ACA	4606

Qy	392	GlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSer	411
Db	4607	CTATTGTTAAACAGTAAATGAAGGAAAAACTTTATATTAAAGATGTCCGAATTGAAGGA	4666
Qy	412	AsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValIlePheProGlu	431
Db	4667	CTTGAACAATAAAGCTAAA-----ACTTTAAAAAAGAGTTTAGCATTAACAGAA	4717
Qy	432	Arg-----GluGlnIleGlnAsn	437
Db	4718	CGTAATTTTATCATGGTTTACTGGAACAGGTGTATTACGTGAAGAATATCTTGAACGT	4777
Qy	438	AspGlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGlu	457
Db	4778	GACTCTATAGCAATC-----TCTGCCTATGCCATCAATCATGCGCTATGTACATATT	4828
Qy	458	SerThrLeuGluPro-----ValIleGluThrValIleGluLeu	469
Db	4829	CAAGTCTGCTCACCTGAAGTAAACATTCAAATGAAAAGGAATTTGTATTACATTTAGAGTA	4888
Qy	470	ThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGln	489
Db	4889	AAAGAAAGTAGCGCTATAAAATAGGAAAATAGACTTTTAAAGGAGATCTTTATTCAGACA	4948
Qy	490	AspLysLeuAsnLeuValAlaIleLysAlaArgHisLeuTyrAspMetProAspArg	509
Db	4949	AATCAACAACCTCTTAAAGTAAACAAAATTTGATGATCATATAAAACATATGACAGATATTTT	5008
Qy	510	ValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAsp	529
Db	5009	TCTCTCTCTGTTATGCAAGATGAT-----GTAAAGCATTTAACAGATTTTTTATTCAGAT	5062
Qy	530	AlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu-- $\bar{\epsilon}$ -----	546
Db	5063	TATGTTTATGCATTTGCTGAAGTAGATCTT---GAACAACCAACAAAATGAACAGATGCA	5119
Qy	547	-----ValIleAspLeuProGluArgThrAlaLeu-----	556
Db	5120	ACAATTGATGTTACTTTCCTTATTGATAAAAAACAAAAGTCTTCTTCGTAGAATAAATT	5179
Qy	557	-----AlaAsnArgLysThrProAlaAspValTyr-----	566
Db	5180	GTTGAAGGAATACTCGTACTAGATAGATAATGTTATCTCCTCGTAATTACGCTTCTGAT	5239
Qy	566	-----	566
Db	5240	GGAGATCTTTTAAATGCTCAACATCTCCGACGCTCTAATGAATGCCTTAACCGGCTTGGC	5299
Qy	567	-----GlnSerLysLysValProLeu	573
Db	5300	TATTTTAACCAAGTAGATACAGATACACTGCCTACAGGAAAGATGATGAAGTTGATCTA	5359
Qy	574	TyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTrpGlySer	593
Db	5360	CTTCTAAAGATT-----CAAGAAGCTCGAACAGGTGCAATCACAGGTGGTGTGGTTTAC	5413
Qy	594	AspThrGlyThrArgLeuValThrLys-----PheGluHisAsnLeuIleAsnArg	610
Db	5414	TCAACACATTTAAATTTGCTGTTTCCAGGAAGTATCTCAAGAAAGAACTATTCGGGGAATA	5473
Qy	611	AspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeu	630
Db	5474	--GGTTATATTTAAGTATTGAAGGTTTTTATTCTAGTAGTCATCTCTCTTGATCTTT	5530
Qy	631	TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr	650
Db	5531	TCTTTTACCAATCCTCGTGTATTATGATACAGAC-----	5563
Qy	651	GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeu-----	664
Db	5564	-----TTTGGCTTTAGTATAATACATTTTATACCTACGAGATGAATGGATGAC	5611

PR	31-MAR-2000; 2000US-0540217.
XX	23-AUG-2000; 2000US-0649167.
XX	(HYSE-) HYSEQ INC.
PA	Dmanac RT, Liu C, Tang YT;
PI	WPT; 2001-639362/73.
DR	P-PSDB; ABG17710.
XX	New isolated polynucleotide and encoded polypeptides, useful in
XX	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity -
XX	
PS	Claim 1; SEQ ID No 17701; 103pp; English.
XX	
CC	The invention relates to isolated polynucleotide (I) and
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	and gene mapping, and in recombinant production of (II). The
CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	a food supplement. (II) and its binding partners are useful in medical
CC	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	disorders involving aberrant protein expression or biological activity.
CC	The polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64197-AAS94564 represent novel human
CC	diagnostic coding sequences of the invention.
CC	Note: the sequence data for this patent did not appear in the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 1425 BP; 299 A; 353 C; 407 G; 366 T; 0 other;
Alignment Scores:	
Pred. NO.:	8.59e-10 Length: 1425
Score:	235.50 Matches: 56
Percent Similarity:	57.26% Conservative: 15
Best Local Similarity:	45.16% Mismatches: 48
Query Match:	4.98% Indels: 5
DB:	23 Gaps: 2
US-09-914-168-2 (1-919) x AAS81897 (1-1425)	
Qy	795 GlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPheAlaGly 814
Db	292 GGCTGGATTGAACCGGTGATTTGCACAAGATACC CGCGGATCTGCGTTTTCTTCGCGGG 351
Qy	815 GlyAspGlnSerIleArgGLYTyrrAlaHisAspSerLeuSerProIleSerAspLysGly 834
Db	352 GCGGACCGCAGTAGTTCTGGTGCTACAAATACAAATCTATCTCCGAAATACGCCACGGT 411
Qy	835 TyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMet 854
Db	412 GACCTGAAAGGGCCCTCGAAGTTGTATACCCGGATCGCTGGAATACCAACGACGTGACC 471
Qy	855 LysaspLeuArgLeuAlavalPheGlyAspIleGlieAsnAlaTyrAspLysGlyPheThr 874
Db	472 GGAATAATGGTGGGCGCGGTGTTTCTCGATAGTGGCGAAGCGGTAAACGATATTCGCGCG 531
Qy	875 AsnAspThrLysIleGlyAlaGlyValGlyValArgTripAlaserProvalGlyGlnVal 894
Db	532 ACCGACTTTAAACCGGTACGCGGGTGGCGTGGGAATCGCGGTTCGGGCGCAATC 591
Qy	895 ArgValasPvalAlaThrGlyVal-----LysGluGluGlyAsnProIleLysLeuHis 912

```

Db 592 AAACGCGATTTCGCGTACCGTCCGGGATAAAGACGACACGGG-----TTACAG 642
Qy 913 PhePheIleGly 916
Db 643 TTTTACATCGGA 654

RESULT 21
ID AAS93087 standard; cDNA: 1425 BP.
AC AAS93087;
XX
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #28891.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX PD 11-OCT-2001.
XX
XX PF 30-MAR-2001; 2001WO-US08631.
XX
XX PR 31-MAR-2000; 2000US-0540217.
XX
XX PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI: 2001-639362/73.
XX
XX P-PSDB: ABG28900.

New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity -
XX
XX Claim 1: SEQ ID No 28891; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 1425 BP; 366 A; 407 C; 353 G; 299 T; 0 other;

Alignment Scores:
Pred. No.: 8.59e-10 Length: 1425
Score: 235.50 Matches: 56
Percent Similarity: 57.26% Conservative: 15
Best Local Similarity: 45.16% Mismatches: 48

```

```

Query Match: 4.98% Indels: 5
DB: 23 Gaps: 2
US-09-914-168-2 (1-919) x AAS93087 (1-1425)

Qy 795 GlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGly 814
||||| : : : : : |||||
Db 1134 GCGTGGATTGAACCGGTGATTTCGACAAAGTACCGCGGATCTCGGTTTCTTGCCTGG 1075

Qy 815 GlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGly 834
||||| : : : : : |||||
Db 1074 GCGACCGCAGTATTGCGTGGCTACAAATACAAATCTATCGTCCGAAATACGCCAACGCT 1015

Qy 835 TyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMet 854
||| ||| ||| ||| ||| ||| ||| |||
Db 1014 GACCTGAAGGGCGCTCGAAGTTGATACCGGATCGCTGGGAATACCAAGTACCAAGTGACC 955

Qy 855 LysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThr 874
||||| : : : : : |||||
Db 954 GGAATATGTTGGGGCGGGTGTTCGATAGTGGCGAAGCGGTAAAGCGATATTCGCGCG 895

Qy 875 AsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyVal 894
: : ||| ||| ||| ||||| ||||| ||||| ||||| ||||| : :
Db 894 AGCGACTTTAAACCGGTACCGGGTGGCGTGGCGTGGGAATCGCGGTTCGGCCCAATC 835

Qy 895 ArgValAspValAlaThrGlyVal-----LysGluGluGlyAsnProIleLysLeuHis 912
: : : : ||| ||| ||| : : : : |||
Db 834 AAACGCGATTTCGCTACCGGTACCGGTACCGGATAACACGACACACGGG-----TTACAG 784

Qy 913 PhePheIleGly 916
||| : : : : |||||
Db 783 TTTTACATCGGA 772

RESULT 22
AAX91724
ID AAX91724 standard; DNA: 2325 BP.
XX
XX AC AAX91724;
XX
XX DT 25-AUG-1999 (first entry)
XX
XX DE Porphyromonas gingivalis protein PG45 encoding DNA.
XX
XX KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis;
XX vaccine; antigenic; ds.
XX
XX OS Porphyromonas gingivalis.
XX
XX PN WO9929870-A1.
XX
XX PD 17-JUN-1999.
XX
XX PF 10-DEC-1998; 98WO-AU01023.
XX
XX PR 04-AUG-1998; 98AU-0005028.
XX
XX PR 10-DEC-1997; 97AU-0000839.
XX
XX PR 31-DEC-1997; 97AU-0001182.
XX
XX PR 30-JAN-1998; 98AU-0001546.
XX
XX PR 10-MAR-1998; 98AU-0002264.
XX
XX PR 09-APR-1998; 98AU-0002911.
XX
XX PR 23-APR-1998; 98AU-0003128.
XX
XX PR 05-MAY-1998; 98AU-0003338.
XX
XX PR 22-MAY-1998; 98AU-0003654.
XX
XX PR 29-JUL-1998; 98AU-0004917.
XX
XX (CSLC-) CSL LTD.
XX
XX PI Agius CT, Barr IC, Hocking DM, Margetts MB, Patterson MA;
XX Ross BC, Rothel LJ, Webb EA;
XX
XX WPI: 1999-385613/32.
XX
XX P-PSDB: AAY34506.

```



```
Qy 790 yGlyIleGlnAlaGlyTyrIleTrpSerAspAsnHisValProTyrArgLeuAr 810
Db 1849 -----AATATGCGAGTGGCACCCCTATAGCGAGCA 1877
Qy 810 gPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeu----- 827
Db 1878 GTTCATGTAGCGGTGCGCAATAGTAGTACAGAGCTTTCACCGTCGCGTACATCGGCCCGCG 1937
Qy 828 -----SerProIleSerAspLysGlyTyrLeu-----ThrGlyGlyGlnValLe 842
Db 1938 ACGGTTCAATCCGGATCCGACAATCAGTATTCTCTATTTGGATCAGGTGGCGGAATTCAA 1997
Qy 842 uAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPh 862
Db 1998 ACTCGAACCCAACTGGGAATATAGAGGCAAGCTTTTCGGGATCTCCACGAGCGGTTT 2057
Qy 862 eGlyAspIleGlyAsnAlaTyr----- 869
Db 2058 CCTCGATCGCGGCAACGTTTGGCTCTTGAGGAGGATTCTTCCCGTCGCGGGGTGCTCT 2117
Qy 870 -----AspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyValG1 884
Db 2118 GTCCGAAGTGGATCGGTGGAGCAATTTCTCGATAGCATCGCTCTCGGCACCGGTGTCGG 2177
Qy 884 yValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThrGlyValLys-- 903
Db 2178 CCTCGCTACGATCGGCATTTCTCGTGGTTCGTGTGATGTCGGCTTCGCTCCACCT 2237
Qy 904 -----GluGluGly-----AsnProIleLy 910
Db 2238 TCCTTACATACGGGTAGAAAGGTACTACAATATCCACGCTTAAAGGATGCCATCGG 2297
Qy 910 sLeuHisPhePheLeGlyThrProPhe 919
Db 2298 TTTCATTTGGCTGCGGCTATCCCTTC 2325

RESULT 23
AAX91597
XX AC AAX91597;
XX DT 25-AUG-1999 (first entry)
XX DE Porphyromonas gingivalis protein PG45 ORF encoding DNA.
XX KW Porphyromonas gingivalis; PG; periodontal disease; gingivitis;
XX OS vaccine; antigenic; ds.
XX OS Porphyromonas gingivalis.
XX PN W09929870-A1.
XX PD 17-JUN-1999.
XX PF 10-DEC-1998; 98WO-AU01023.
XX PR 04-AUG-1998; 98AU-0005028.
XX PR 10-DEC-1997; 97AU-0000839.
XX PR 31-DEC-1997; 97AU-0001182.
XX PR 30-JAN-1998; 98AU-0001546.
XX PR 10-MAR-1998; 98AU-0002264.
XX PR 09-APR-1998; 98AU-0002911.
XX PR 23-APR-1998; 98AU-0003128.
XX PR 05-MAY-1998; 98AU-0003358.
XX PR 22-MAY-1998; 98AU-0003654.
XX PR 29-JUL-1998; 98AU-0004917.
XX (CSLC-) CSL LTD.
XX PA Agilus CT, Barr IG, Hocking DM, Margetts MB, Patterson MA;
XX PI Ross BC, RotheL LJ, Webb EA;
```

```
XX WPI; 1999-385613/32.
DR P-PSDB; AAY34379.
XX Antigenic Porphyromonas gingivalis peptides for preventing
PT gingivitis
```

XX Claim 12; Page 136-137; 588pp; English.

XX AAX91536 to AAX91801 encode two hundred and sixty six antigenic Porphyromonas gingivalis (PG) polypeptide sequences given in AAY34318 to AAY34583. AAX91802 to AAX91989 represent PCR primers used in the isolation of the PG polypeptides. The PG polypeptides have antibacterial activity with a vaccine mechanism of action. The PG polypeptides can be used as vaccines especially against Porphyromonas gingivalis. Probes can be used to detect Porphyromonas gingivalis in standard hybridisation assays. Porphyromonas gingivalis is involved in periodontal disease especially gingivitis.

XX Sequence 2409 BP; 603 A; 657 C; 569 G; 580 T; 0 other;

Alignment Scores:
Pred. No.: 5.79e-09 Length: 2409
Score: 229.00 Matches: 161
Percent Similarity: 32.99% Conservative: 93
Best Local Similarity: 20.91% Mismatches: 264
Query Match: 4.84% Indels: 253
DB: 20 Gaps: 35

US-09-914-168-2 (1-919) x AAX91597 (1-2409)

```
Qy 300 ProLeuLeuIleGlyAspValPheHisGlyLysTyrGluThrLysLysAsnLeuIle 319
Db 409 CCGTTTTTCATCATCAGTCAGGTC-----AAATCCGATAGCGGCTAAGGTGGCG 456
Qy 320 GluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAspArgSerValAsp 339
Db 457 ACGAATCTCCGCGAAGACGCGGTACTTCGATGCTAA---GTAAAGACAGTGTGACC 513
Qy 340 ValIleLeuProAspAsn---ThrAlaAspValSerLeuIleTyrAspThrGlyThrGln 358
Db 514 ACTCTGAAAAAGGACTCGCTCAAGCCCAAAATCTCTATACGGTGGATATGCGCTCTCT 573
Qy 359 TyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThr 378
Db 574 TATCATTCAGACATCATTCCTCCCTTACCAGTC-----AGCACT 612
Qy 379 AspProAspLysLeuProValLysArgGluLeuLeuGluInLeuLeuThrValAsnMet 398
Db 613 TTCCCGACAGCATTTCTGGCTTACAGGCGACTCGCTCT-----TTGATCAGGAAA 663
Qy 399 GlyGluAlaTyrAsnLeu-----GlnAlaValArgAlaLeuSer 411
Db 664 GGAGACCAAGTTCAATTTGGCAAGCTGCACGAGAGCGTCAGACCATCAGTCCCTGCTG 723
Qy 412 AsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGlu 431
Db 724 AGAGAC-----AATGGTTACTACTTCTCCGCCACAGGATATATCTACGAAGCC 774
Qy 432 ArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArgThrGlu 451
Db 775 GATACCTCTCTCGTAAGAGGTGCGCTATGCTCGGAGCCCAAGCTCTCGGAAGATACCTCA 834
Qy 452 ProAla-----GlnValAspGluSerThrLeu----- 460
Db 835 CCCCAAGCCATGCGCCCGTGGAGGATAGGAAACGACAGCAGTCTCTCGGAANTGAAC 894
Qy 461 ---GluProValIleGluThrValGluLeuThrAsp---GlyIleLeu----- 474
Db 895 GGAGAAAGCCCGACAGACTCGCTCGAGTGGAGGATATGAAGTCTTTTACTATCGTAAA 954
Qy 475 MetAspIleSerPro-----IleGluPheSerAlaSerAsnLeuIle 488
```

```
Db 955 ATGCCGGTTCGCCCAAGATTTTGGCCAAACGCTTTCGTTCTTCCTCCGGCAATCTGTAT 1014
Qy 489 GlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp 508
Db 1015 CGGCAGAA----- 1023
Qy 509 ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSer 528
Db 1024 -----GAGCATGAGACGACACGCAAAATCCTTGGCTCGTTGGGA 1062
Qy 529 AspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIle 548
Db 1063 ---GCCCTTCCTCGTATCGATCTCAATTTTGGCAACGATTCATTTCCGGCGCTTTTG 1119
Qy 549 AspLeuProGluArgThrAlaLeu-----AlaAsnArgLysThrProAla 563
Db 1120 GATCTGGCACTGCTAACCCACCTCGACNAACCTTGGGATGATCATTTAGAGACC----- 1173
Qy 564 AspValTyrGlnSerLysLys-----ValProLeuTyrValPheValAla 578
Db 1174 ---TTGTTTCAGAGCAAAAGCAATGACTTTCATCGGTCGCCGACTCAATTTTGCTTGTCT 1230
Qy 579 SerAspLysProArgAspGlyGlnIleGlyLeuGlyTyr-----GlySer----- 593
Db 1231 CGGCGCAATGATTCGGCGGAGGAGAAATCTTTCTTGGAAATATCGGTGGATCGTATGAG 1290
Qy 594 ---AspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGly 612
Db 1291 TGGGAGACCGGCAATCGT---CCGCAAAATAGCAGCAATCGGCTCATATAAATTCG 1347
Qy 613 TyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeuTyrAla 632
Db 1348 TACAACATGAATACGGCGGTGAACCTCTCGTTTCCCTCGATGTATTTTCCCGGCTGCTG 1407
Qy 633 ThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThr----- 647
Db 1408 GATAAATACTATTACTACCCACGACTACGACTTTTTCAGGCTTCGCCACCGCGCTGAAC 1467
Qy 647 ----- 647
Db 1468 AGGCACACTACTTTAGCATGTACTTTCGGCTTTTCGACCACTACGAATTTTCAGCCC 1527
Qy 648 -----LeuGlyTyrGlnGlnGluValPheGlyHis 657
Db 1528 TCCAAGGAACACCGCATGCTATTTTCCCGCTCAAGCTCAACTACAACCTCTCGGGCAT 1587
Qy 658 SerThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHisGluIleSer 674
Db 1588 CAGACAGAAACTTTCAGGCGCATACGGCGAACAATCCGCCCTGCT-GCTCAGCCTTCA 1646
Qy 675 ArgSerIle---IleGlnAsnGlyGlyTrpAsnArgThrTyrSerLeuArgTyrArg--- 692
Db 1647 GAGTCAGTCTCTCAATATGGGTATATCTATACGTTCAACAATCCGTTTCAGAGAA 1706
Qy 693 -----LeuAspLysLeuLysThrGlnAlaProGluThrTrpGlnAsp 707
Db 1707 AAGTCCTCATCATCTTTTGGATGCAATTCGGACT----- 1739
Qy 708 LeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGly----- 725
Db 1740 -----ATCCGAGGAGGCAATCTCTGATCTGATCTGATCTGCGGACCGG 1784
Qy 726 -----ValAlaValHis-LysTh 731
Db 1785 CAAGAAGTACAGCAGACCAAGAATTTTCGTCGGGTCCTCTCTCTCATCAAAAGC 1844
Qy 731 rValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluVa 751
Db 1845 CACGGAGAA-----CTGCGCTATTCTTACCATT 1874
Qy 751 lGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAla---ArgAlaGlyIleSe 770
Db 1875 A-----GACCGCAATCAGTCACTGGCAACCCGTTTCGGGACAGG 1913
```

```
Qy 770 rGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGl 790
Db 1914 CGTGATATATAGCTATGCG----- 1932
Qy 790 yGlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArgLeuAr 810
Db 1933 -----AATATGCGAGTGGCACCTATAGCAGCA 1961
Qy 810 qPhePheAlaGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeu----- 827
Db 1962 GTTCTATAGCGGTGCCAATAGTATCAGAGCTTTACCGCTCCGTAGCATCGGCCCGG 2021
Qy 828 -----SerProIleSerAspLysGlyTyrLeu-----ThrGlyGlyGlnValIle 842
Db 2022 ACGGTTCAATCCGATTCGACAAATCAGTATTCCTATTGGATCAGGTGGCGAATTCAA 2081
Qy 842 uAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPh 862
Db 2082 ACTCGAAGCCCAACGTGGAATATAGAGGCAAGCTTTTCGGGATCTCCACGACGCGTTT 2141
Qy 862 eGlyAspIleGlyAsnAlaTyr----- 869
Db 2142 CCTCGATCGGCAACGTTTGGCTCTTCAGGAGGATTCCTCCCGCTCCGGCGGTGCTCT 2201
Qy 870 -----AspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGl 884
Db 2202 GTCGGAAGTGGGATCGGTGAGCAATTTCTCTGAATAGCATCGCTCTCGCACCGGTGTCG 2261
Qy 884 yValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThrGlyValLys--- 903
Db 2262 CTTTCGCTACGATCGGCATTTTCGTTGTCGATGTCGCTTCGCTTCGCTTCACCT 2321
Qy 904 -----GluGluGly-----AsnProIleLey 910
Db 2322 TCTTTACAATACGGGTAGAAGGTTACTACAATATCCACCGCTTAAGGATGCCATCGG 2381
Qy 910 sLeuHisPheIleGlyThrProPhe 919
Db 2382 TTTCCATTTGGCTGTCGGCTATCCCTTC 2409
RESULT 24
AAA48507
ID AAA48507 standard; DNA; 1830 BP.
XX
AC AAA48507;
XX
XX 03-JAN-2001 (first entry)
XX
XX Neisseria meningitidis BASB040 putative coding sequence #1.
XX
XX BASB040; bacterial disease; respiratory tract infection; bacteraemia;
XX
XX meningitis; cancer; autoimmune disease; ds.
XX
XX Neisseria meningitidis.
XX
XX key Location/Qualifiers
XX CDS 1..1830
XX /*tag= a
XX /product= "BASB040"
XX
XX WO200034480-A1.
XX
XX 15-JUN-2000.
XX
XX 02-DEC-1999; 99WO-EP09560.
XX
XX 07-DEC-1998; 98GB-0026886.
XX
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Ruelle J;
XX
```


QY 899 AlaThrGlyValLysGluCluGlyAsnProIleLysLeuHisPhePheIleGlyThrPro 918
 ||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 1771 GCCTACGGGCACAGCGATAAG-----AAATCCGCTGGCATCATCAGCTTGGGAACACGC 1824

QY 919 Phe 919
 |||

Db 1825 TTC 1827

RESULT 25

AAA48508

ID AAA48508 standard; DNA; 1830 BP.

XX AC AAA48508;

XX DT 03-JAN-2001 (first entry)

XX DE Neisseria meningitidis BASB040 putative coding sequence #2.

XX KW BASB040; bacterial disease; respiratory tract infection; bacteraemia;
 XX KW meningitis; cancer; autoimmune disease; ds.

XX OS Neisseria meningitidis.

XX FH Key Location/Qualifiers
 XX CDS 1..1830
 XX FT /*tag= a
 XX FT /product= "BASB040"

XX PN WO200034480-A1.

XX PD 15-JUN-2000.

XX PF 02-DEC-1999; 99WO-EP09560.

XX PR 07-DEC-1998; 98GB-0026886.

XX PA (SMK) SMITHKLINE BEECHAM BIOLOGICALS.

XX PI Ruelle J;

XX DR WPI; 2000-423426/36.

XX DR P-PSDB; AAY99624.

XX PT Novel BASB040 polypeptides of Neisseria meningitidis useful for
 PT diagnostic, prophylactic and therapeutic purposes against microbial
 PT diseases comprise a specific amino acid sequence

XX PS Claim 11; Page 60; 98pp; English.

XX CC The present sequence is a putative version of the Neisseria
 CC meningitidis strain ATCC13090 BASB040 coding sequence. The protein
 CC produced from this gene is similar to the p15 outer membrane protein of
 CC the bacterium. The gene, its protein, antibodies, antagonists and
 CC agonists can be used to diagnose and treat bacterial diseases such as
 CC those leading to upper respiratory tract infections, bacteraemia and
 CC meningitis. In addition, they can be used in vaccines for use against
 CC cancer and autoimmune diseases.

XX SQ Sequence 1830 BP; 439 A; 604 C; 461 G; 326 T; 0 other;

Alignment Scores:

Pred. No.: 9.4e-09 Length: 1830
 Score: 224.50 Matches: 135
 Percent Similarity: 34.33% Conservative: 72
 Best Local Similarity: 22.39% Mismatches: 218
 Query Match: 4.75% Indels: 178
 DB: 21 Gaps: 19

US-09-914-168-2 (1-919) x AAA48508 (1-1830)

QY 462 ProValIleGluThrValCluLeuThrAspGlyIleLeuMetAspIleSerProIleGlu 481
 ||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 124 CCCGACACGAATCAGTTAAATTAACCCAAATTCCTCCGCGTCATCGACACGCGGAT 183

QY 482 PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHis 501
 : : : : : : : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 184 AGTGAATCAAGATATGCTCGAAGAACACCTGCGCCTCATCAGCAGCAGCAGGAGAA 243
 QY 502 LeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArg 521
 : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 244 GTATTGGACAAAGAACAGACAGCGGCTTCTTCGCGAAGAACACCGGACACCGTTAAAC 303
 QY 522 SerIle-----LeuGlyArgIleSerAspAlaValSerAla----- 533
 : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 304 ATGCTCCGACGAAGGCTATTTCACAGCAAAAGTCAGCCTGACGGAAAAAGACGAGCT 363
 QY 534 -----ValAlaArgAlaIle 538
 : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 364 TATACGGTACATCATCACCGCGCGCCACCAAAATCCCAACCTGCTGTCGCATC 423
 QY 539 LeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsn 558
 ||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 424 CTCGGGACATCTTTTCAGACGGCAACCTCGCCGAATACATACCGCAACGCGTGGAAAC 483
 QY 559 ArgLysThrProAla-----AspValTyrGlnSerLysLys----- 570
 : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 484 TGGCAGCAGCCGCTAGGCACTGATTCGATCAGGACAGTTGGGAAACACGCAAACTTCC 543
 QY 570 ----- 570
 Db 544 GTCTCGCGCGGTAAACGGCAAGGCTACCGCTTGCCAAAGCTCGGCAACACCCGGCG 603
 QY 571 -----ValProLeuTyrValPheValAlaSerAspLysPro 582
 ||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 604 GCCGTCAACCCCGATACCCACCGCTGATTTGAAGTCGTCGTGGACAGCGGCCGCC 663
 QY 583 ArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArg----- 598
 ||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 664 -----ATCGCTTCGGCGACTTTGAAATCACCGGCACACAGCGTTACCCCGAA 711
 QY 599 -----LeuVal 600
 : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 712 CAATCGTCTCGGCGCTTCCAAACGGGCGACGCCCTACGACCTCGACCTGCTG 771
 QY 601 ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeu--- 619
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 772 CTCGACTTCCCAACGACACTC---GAACAAACGGGCATTATTCGGCGCGCTCCGACAA 828
 QY 620 -----ArgLeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysPro 635
 ||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 829 GCCGACTTCGACCGTCTCCAAAGCGACCGCTCCCGCTCAAGTCAGCTAACCGAGGTC 888
 QY 636 LeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPhe 655
 ||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 889 AAACGCCACAAGCTCGAAACCGGCATCCGCTCGATTCCGAATAC----- 933
 QY 656 GlyHisSerThrAsnGly-----PheAsp-LeuSerThrArgThrLeuGluHisG1 672
 ||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 934 -----GGTTTGGGCGCAAAATCGCTACGACTATTACAACCTCTTCAACAAAGGCTAT 987
 QY 672 uIleSerArgSerIleIleGlnAsnGlyTyrAsnArgThrTyrSerLeuArgTyrAr 692
 : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 988 ATCGCTCGTCTG-----TCTGGGATATGGCAATACG----- 1021
 QY 692 gLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPh 712
 ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 1022 -----AAACACAGCTTGGCGCGCATCAGCAGCGCGCAACTATTCGGGC 1068
 QY 712 eValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrVa 732
 ||| ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 1069 AACTACTGGA-----CAAGCAACGTTTCTTACAACCGTTTCGACCAACCAACCC-- 1117
 QY 732 lAlaAspAsnLeuValAsnPro----- 740
 : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
 Db 1118 -TCGAAAAACGCCCTTCTCCGCGGCATCTGTGTGTCGGCAGCGCGCGGCATCGAT 1176


```

Db 454 -----CTTTCAGACGGCAACCTCCCGCAATACTACCGCAACGCGCTGGAAACTGG 504
Qy 478 ---SerProIle-----GluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeu 494
Db 505 CAGCAGCGGTAGGACAGTTCATCAGGACAGTTCGGAAACAGCAAACTTCGGTC 564
Qy 495 ValAlaAlaLysAlaArgHisLeuTyrAspMetPro-----AspAspArgValLeuAla 512
Db 565 CTCGCGCGGTAAACGCAAGCCTACCGCTTGCAAGCTCGGCAACACCGCGGCGCC 624
Qy 513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532
Db 625 GTCAACCCCGATGACGCC-----ACCGCGATTGAAC 657
Qy 533 AlaValAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
Db 658 GTCGTCGTGGACAGCGCGGCCCA-----TCGCTTCGGGCACTTTCGAATATACCGC 711
Qy 552 uArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValPr 572
Db 712 ACGACGCTTACCCGGAACAATCTCTCCGGCT----- 745
Qy 572 oLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTrpGl 592
Db 746 -----TCGCGCGCTTCCACCGGCGCGC----- 769
Qy 592 ySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGl 612
Db 770 ---CCTAGCACCTCGAC-CTGCTCTCGACTTCCACAGCGGCTC---GAACAACAGG 821
Qy 612 yTyrGlnAlaGlyAlaGluLeu-----ArgLeuSerGluAspLysLysGl 627
Db 822 GCATTATTCCGCGCGGTCCGTACAACGCGACTTTCACGCGCTCCCAAGCGCGCTGCC 881
Qy 627 yValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaTh 647
Db 882 CGTCAAGTCACGTAACAGGTCGTAACAGCGCAAGCTCGAAACCGCGCATCCGCTCGA 941
Qy 647 rLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly----- 661
Db 942 TTCGGAATAC-----GGTTGGCGCGCAAAATCGCTACGACTA 980
Qy 662 -PheAspLeu-SerThrArgThrLeuGluHisGluIleSerArgSerIleLeuGlnAsnG 681
Db 981 TTACAACCTCTTCAACAAGGCTATA-----TCGGTTCGGTCTGCTG 1022
Qy 681 yGlyTyrAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaP 701
Db 1023 GATATGGACAATAACG-----MetArgGlyTyrArgGlnArgTyrSerL 749
Qy 701 roProGluThrTrpGlnAspLeuProValAspPheValasnGlyLysProSerGlnGluA 721
Db 1053 CGCGCGCATCAGCCAGCGCGCAACTATCGGGCACTACTGGA-----CAAGCAA 1103
Qy 721 laLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnPro- 739
Db 1104 CGTTTCTTACACCGCTTCGACCAACCAAAACC---TCGAAAAACGCGCTTCTCGCGGG 1160
Qy 740 -----MetArgGlyTyrArgGlnArgTyrSerL 749
Db 1161 CATCTGGTATGTCGCGACCGCGCGGATCGATCCAGCTGGGGCGGAGTTTCTCGC 1220
Qy 749 euGluValGlySerSer---GlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGl 768
Db 1221 AGAAGCGCGGAAATCCCGCGC-----TCGGATATCGATTGGGCAACAGCCGCCAC 1274
Qy 768 yIle-----SerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAl 785
Db 1275 GATGCTGACCGCTCTTGGAAACGCGCAGCTGCTCAACACGCTGCTGCATCCGGAACAGG 1334
Qy 785 aHisGlnMetThrGlyGlyIle----- 792
Db 1335 CCATTACCTCGACGCAAAATCGGTACGACTTTGGCGCAATTCCTGTCTCCTCCACGCGCT 1394

```

```

Qy 792 ----- 792
Db 1395 GATCCGCACCTCTGCCGCGCAGGTATTCTTTCAGCCCGGAAACAAAACTCGGCAC 1454
Qy 793 -----GlnAlaGlyTyr---IleTrpSerAspAsnPheAsnHisValPr 806
Db 1455 GTTCATCATACGCGACAAGCGGTTACACCGTTGCGCGCACAATGCCAAC---GTTCC 1511
Qy 806 oTyrArgLeuArgPhePheAlaGlyClyAspGlnSerIleArgGlyTyrAlaHisaspSe 826
Db 1512 TTCAGGGCTGATGTTCCGACGCGCGCGCTTCCGTCGCGGTTCAGAACTCGACAG 1571
Qy 826 rLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyTh 846
Db 1572 CATCGGGCTTCCGCGCGCAACGGATCGCTCTCCCGAACCGCCCTCTTGTGGCGAG 1631
Qy 846 rAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGl 866
Db 1632 CCGTAATACCAACTGCGGCTTTACGCGCACCGCTTTCCGCGCGGTGTTCCACGATATGG 1691
Qy 866 yAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValAr 886
Db 1692 CGACCGCGCGCAATTCAAACGATGATGAACGCTGAACACGCTCGGACTCGGCTGCG 1751
Qy 886 gTrpAlaSerProValGlyGlnValArgValAspValAlaThrGlyValLysGluGluGl 906
Db 1752 CTGGTTACGCGCGCTCGCGCGCTTTTCCTTCGACATCGCTACGCGCACGCAACAAG-- 1809
Qy 906 yAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
Db 1810 ----AAATCCGCTGGCACATCAGCTTGGAAACGCGCTTC 1845

RESULT 27
AAA48509
ID AAA48509 standard; DNA; 1764 BP.
AC AAA48509;
XX
XX
XX 03-JAN-2001 (first entry)
XX
XX Neisseria meningitidis BASB040 putative coding sequence.
XX
XX BASB040; bacterial disease; respiratory tract infection; bacteraemia;
XX meningitis; cancer; autoimmune disease; ds.
XX
XX Neisseria meningitidis.
XX
XX Key Location/Qualifiers
XX CDS 1..1764
XX FT /*tag= a
XX FT /product= "BASB040"
XX FT /partial
XX
XX WO200034480-A1.
XX
XX 15-JUN-2000.
XX
XX 02-DEC-1999; 99WO-EP09560.
XX
XX 07-DEC-1998; 98GB-0026886.
XX
XX (SMIK ) SMITHKLINE BEECHAM BIOLOGICALS.
XX
XX Ruelle J;
XX
XX WPI; 2000-423426/36.
XX DR P-PSDB; AAY99625.
XX
XX Novel BASB040 polypeptides of Neisseria meningitidis useful for
XX diagnostic, prophylactic and therapeutic purposes against microbial
XX diseases comprise a specific amino acid sequence
XX

```

PS Claim 11; Page 61; 98pp; English.

CC The present sequence is a putative version of the *Neisseria meningitidis* strain H44/76 BASB040 coding sequence. The protein CC produced from this gene is similar to the D15 outer membrane protein of CC the bacterium. The gene, its protein, antibodies, antagonists and CC agonists can be used to diagnose and treat bacterial diseases such as CC those leading to upper respiratory tract infections, bacteraemia and CC meningitis. In addition, they can be used in vaccines for use against CC cancer and autoimmune diseases.

XX Sequence 1764 BP; 424 A; 581 C; 451 G; 308 T; 0 other;

Alignment Scores:

Pred. No.:	4.02e-08	Length:	1764
Score:	216.50	Matches:	136
Percent Similarity:	34.87%	Conservative:	76
Best Local Similarity:	22.37%	Mismatches:	208
Query Match:	4.58%	Indels:	188
DB:	21	Gaps:	21

US-09-914-168-2 (1-919) x AAA48509 (1-1764)

Qy	462	ProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGlu	481
Db	58	CCCGACACCGAATCAGTCMAATTAACCCCAAAATCCCGCTCCTCATCGACACCGAGAC	117
Qy	482	PheSerAlaSerAsnLeuIleGlnAspLysLeuValAlaAlaLysAlaArgHis	501
Db	118	AGTGAATCAAGATATGTCGAAGAACACCTCGCGTCTCATCGACGACGAGGAAGAA	177
Qy	502	LeuTyrAspMetProAspAspArgValIleAlaIleAsnHisAspAspGlyValAsnArg	521
Db	178	GTATTGACAGGACAGACGGGCTTCCTCGCCGAAGACGGCGGACACAGTTTAAACG	237
Qy	522	SerIle-----LeuGlyArgIleSerAspAlaValSerAla-----	533
Db	238	ATGCTCGCAGCAAGAGCTATTTCAGCAGCAAAAGTCAGCTGACGGAAAGACGGAGCT	297
Qy	534	-----ValAlaArgAlaIle	538
Db	298	TATAGGTACACATCACACCGGCGCCGCCAACCAAAATCGCCACGCTCGCGCTGCGCATC	357
Qy	539	LeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsn	558
Db	358	CTCGGCACATCTTTCAGACGGCAACTCGCCGAATACTACCCCAACGCGCTGGAAC	417
Qy	559	ArgLysThrProAla-----AspValTyrGlnSerLys-----	570
Db	418	TGGCAGCAGCGGTAGCAGCGATTTCGATCAGACAGATTGGGAACACGCAAACTCC	477
Qy	570	-----	570
Db	478	GTCTCGGCGGTAGCGCAAGCCTACCCGCTTGCCAAAGCTCGGCAATACGACGGCG	537
Qy	571	-----ValProLeuTyrValPheValAlaSerAspLysPro	582
Db	538	GCGCTCAACCCGATACCGCCACCGCGGATTTGAAGCTGCTGCGGACACGCGCGCCC	597
Qy	583	ArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrArg-----	598
Db	598	-----ATCGCTTCGCGGACTTTGAATACCCGGCACACAGCGCTTACCCCGAA	645
Qy	599	-----LeuVal	600
Db	646	CAAAATGCTCTCCGCGCTTGCGGCTTTCCAGCCGCGGTATGCGGTACGACCTCGACCTGCTG	705
Qy	601	ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeu---	619
Db	706	CTCGACTTCAACAGCGGCTC-----GAACAAACGGGCTATTATTCGCGCGGCTCCGTACAA	762
Qy	620	-----ArgLeuSerGluAspLysGlyValLysLeuTyrAlaThrLysPro	635

Db	763	GCCGACTTCGACCGCTCCAAAGGCGACCGCGTCCCGCTCAAAAGTCAGCGTANCCAGGTC	822
Qy	636	LeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPhe	655
Db	823	AAAGCCCAAAATCGAAACCGGCATCGCTCGATTCCGAATAC-----	867
Qy	656	GlyHisSerThrAsnGly-----PheAspLeu-SerThrArgThrLe	669
Db	868	-----GGTTTGGCGCGCAAAATCGCTACGACTATTACAACCTCTTCAACAAGGCTAT	921
Qy	669	uGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThrTyrSerLe	689
Db	922	A-----TCGGTTCGGTCTCGGTATATGGACAATAACG-----	955
Qy	689	uArgTyrArgLeuAspLysLeuLys-----	697
Db	956	AAACACGCTTCGCGCGGCATCAGCCAGCGCGCAACTATCGGGGCACTACTGGACA	1014
Qy	698	-----ThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPh	712
Db	1015	AGCAACGTTTCTTACACCGTTTCGACCACCCAAACCTCGAAAACCGCGCTTCTCCGCG	1074
Qy	712	evalAsnGly-----LysProSe	718
Db	1075	GGCTCTGGTATGTCGGGACCGCGGGCATCGATGCCAGCTGGGGCGGGAATTTCTC	1134
Qy	718	rGlnGluAla--LeuLeuAlaGlyValAlaVal-----HisLysThr---	731
Db	1135	GCAGAAGCGCGAAATCCCGCGCTCGCTCGATTGGGCAACAGCCAGCGCATG	1194
Qy	732	-----ValAlaAspAsnLeuValAsnProMetArgGlyTyr	743
Db	1195	CTGACCGCTTTCGAAACGCGCATCTCAACACCTGTCATCCGCAAAACGCGCAT	1254
Qy	744	-----ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuVal	757
Db	1255	TACCTCGACGCCAAATATCGGTACGACTTTGGGCACATTCTCTCCACCGCGCTGATC	1314
Qy	758	SerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp	777
Db	1315	CGCACCTCT-----GCCGTCGAGT-----	1335
Qy	778	AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIle-----Gln	793
Db	1336	-----TATTTCTTCACGCGCGCAAAACAAACAACTCGGACGTTTCATCATACGCGGACA	1389
Qy	794	AlaGlyTyrIleTrpSerAspAsnHisValProTyrArgLeuArgPhePheAla	813
Db	1390	GCGGTTTACCGTTGCGCGGACAAATCCCGACGTTCTCTCAGGCTGATGTTCCGCGAGC	1449
Qy	814	GlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLys	833
Db	1450	GCGCGCGCTTTCGCTGCGCGGTACGAACCTCGACAGCATCGGACCTTGGCGCGCGAAC	1509
Qy	834	GlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPhe	853
Db	1510	GGATCGCTCTGCGCGAACCGCGCTCTCTGCTGGGACGCTTGAATACCAACTGCCGTTT	1569
Qy	854	MetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPhe	873
Db	1570	ACGGCACCTTTCCGCGCGGTGTTCCAGATATGGCGATGCCCGCGCAATTTCCAAA	1629
Qy	874	ThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyGln	893
Db	1630	CGTATGAAGCTGAACACGCTTCGGGACGCTGGGCTGCGCTGTTACGCCCGCTCCGCGC	1689
Qy	894	ValArgValAspValAlaThrGlyValLysGluGlyAsnProIleLysLeuHisPhe	913
Db	1690	TTTTTCTTCACATCGCTACGGGACGACGAGGATAAG-----AAAATCCGCTGGCACATC	1743
Qy	914	PheIleGlyThrProPhe	919
Db	1744	AGCTTGGGAACACGCTTC	1761

RESULT 28
AAZ53623
ID AAZ53623 standard; DNA; 1848 BP.
XX AC AAZ53623;
XX 21-MAR-2000 (first entry)
XX DE Neisseria meningitidis ORF 286 partial DNA sequence SEQ ID NO:1195.
XX DE Neisseria meningitidis; Neisseria gonorrhoeae; antigen; vaccine;
KW antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW antibacterial; gene therapy; ds.
XX OS Neisseria meningitidis.
XX PN WO957280-A2.
XX PD 11-NOV-1999.
XX PF 30-APR-1999; 99WO-US09346.
XX PR 01-MAY-1998; 98US-0083758.
XX PR 31-JUL-1998; 98US-0094869.
XX PR 02-SEP-1998; 98US-0098994.
XX PR 02-SEP-1998; 98US-0099062.
XX PR 09-OCT-1998; 98US-0103749.
XX PR 09-OCT-1998; 98US-0103794.
XX PR 09-OCT-1998; 98US-0103796.
XX PR 25-FEB-1999; 99US-0121528.
XX {CHIR } CHIRON CORP.
PA {GENO- } INST GENOMIC RES.
PA Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI Petersen J, Pizsa M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI Tettelin H, Venter JC;
XX WPI: 2000-062150/05.
DR P-PSDB; AAY74861.
XX
XX Novel Neisserial polypeptides predicted to be useful antigens for
PT vaccines and diagnostics -
XX Claim 7: Page 667; 1453pp; English.
XX
XX AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC represent novel Neisseria meningitis and N. gonorrhoeae polynucleotides
CC and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ55473 represent
CC PCR primers used in the exemplification of the present invention. The
CC polypeptides, the polynucleotides, antibodies and compositions of
CC the invention can be used as vaccines, as diagnostic reagents, and as
CC immunogenic compositions. The polypeptides can be used in the
CC manufacture of medicaments for treating or preventing infection due to
CC Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC presence of Neisseria bacteria, or to raise antibodies. They may also
CC be used to screen for agonists or antagonists, which may themselves
CC have use as antibacterial agents. The polynucleotides of the invention
CC may also be used in gene therapy protocols.
XX
XX SQ Sequence 1848 BP; 439 A; 613 C; 467 G; 329 T; 0 other;

Alignment Scores:
Pred. No.: 4,27e-08 Length: 1848
Score: 216.50 Matches: 136
Percent Similarity: 34.87% Conservative: 76
Best Local Similarity: 22.37% Mismatches: 208
Query Match: Indels: 188
DB: Gaps: 21

US-09-914-168-2 (1-919) x AAZ53623 (1-1848)

```
Db 1159 GCGGCTCGTATGTCGCGACGCGCGCGGCGATCGATGCCAGCTGGGCGCGGAATTTCTC 1218
      :::::
Qy 718 rGlnGluAla--LeuLeuAlaGlyValAlaVal-----HisLysThr--- 731
      |||:::||||
Db 1219 CGAAGAGCGGAAATCCCGGCTCGCTGCTGATTTGGCAACAGCAGCCACCGATG 1278
      :::::||||
Qy 732 -----ValAlaAspAsnLeuValAlaValProMetArgGlyTyr 743
      :::::||||:||||
Db 1279 CTGACCGCTCTTGAAACGCGAGCTGCTCAACAGCTGCTCATCCGAAACGCGCAT 1338
      :::::
Qy 744 -----ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuVal 757
      :::::||||:||||
Db 1339 TACCTCGACGCAAAATCGGTACGACTTTGGGCACATTCCTGCTCCACCGCGTGTATC 1398
      :::::
Qy 758 SerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp 777
      :::::
Db 1399 GGCACCTCT-----GCCGCTGCAGT----- 1419
Qy 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIle-----Gln 793
      :::::
Db 1420 -----TATTTCTTCACGCGCGCAACAACTCGGCAGCTTCATCATACGCGGACAA 1473
      :::::
Qy 794 AlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAla 813
      :::::
Db 1474 GCGGTTACACCGTTCGCGCGGAGTACGAACTCGACAGCATCGGACTTGGCGGCCGGAAC 1593
      :::::
Qy 814 GlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLys 833
      :::::
Db 1534 GCGGCGCGCTTCCTCGCGCGGTAGCAACTCGACAGCATCGGACTTGGCGGCCGGAAC 1593
      :::::
Qy 834 GlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPhe 853
      :::::
Db 1594 GATCGGTCTCGCGCGCGCGCTCTGCTGGCGGAGCTGCGGAAATACCAACTGCCGTTT 1653
      :::::
Qy 854 MetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPhe 873
      :::::
Db 1654 ACGGCACCTTTCCGGCGCGGTGTTCCACGATATGGCGGATCGCGCGCAATTTCAA 1713
      :::::
Qy 874 ThrAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyGln 893
      :::::
Db 1714 CGTATGAAGCTGAACACGCGTTCGGGACTGGCGCTGGCTGCTCAGCGCGCTTGGCGCG 1773
      :::::
Qy 894 ValArgValAspValAlaThrGlyValLysGluGlyAsnProIleLysLeuHisPhe 913
      :::::
Db 1774 TTTTCCTTCGACATCGCTACGGGCACAGCGATGAAG-----AAATCCGCTGGCACATC 1827
      :::::
Qy 914 PheIleGlyThrProPhe 919
      :::::
Db 1828 AGCTTGGGAACGCGCTTC 1845
      :::::
RESULT 29
AA81473
ID AAA81473 standard; DNA; 92934 BP.
XX
AC AAA81473;
XX
AC
XX
DT 04-DEC-2000 (first entry)
XX
DE N. meningitidis partial DNA sequence gnm_21 SEQ ID NO:21.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
KW Meningococcus B; MenB; ds.
XX
OS Neisseria meningitidis.
XX
PN WO200022430-A2.
XX
PD 20-APR-2000.
XX
PF 08-OCT-1999; 99WO-0523573.
XX
XX
```

```
PR 09-OCT-1998; 98US-0103794.
PR 30-APR-1999; 99US-0132068.
XX (CHIR ) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Pizza M;
XX
XX WPI: 2000-318079/27.
XX
XX Isolated nucleotide sequences of Neisseria meningitidis which can be
XX used in the diagnosis and treatment of N. meningitidis infection and
XX other Neisserial infections, for example, N.gonorrhoea.
XX
XX Claim 7; Page 471-498; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic
XX proteins from Neisseria genomic sequences. AAA81453 to AAA82414
XX represent specifically claimed Neisseria meningitidis genomic DNA
XX sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
XX Neisseria DNA sequences and their corresponding probes; AAA81254 to
XX AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
XX isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
XX AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
XX sequences, which are all used in the exemplification of the present
XX invention. The nucleic acid sequences, protein sequences, and antibodies
XX against them, can be used in the manufacture of a composition. The
XX composition can be used as a medicament (or in the manufacture of a
XX medicament) for treating, preventing or diagnosing infection due to
XX Neisserial bacteria. For example, some of the identified proteins could
XX be components of vaccines against Meningococcus B; against all serotypes;
XX and/or against all pathogenic Neisseriae. Identification of sequences
XX from the bacterium will also facilitate production of biological probes,
XX particularly organism-specific probes. Attempts to make efficacious
XX Meningococcus B vaccines have failed mainly due to antigen tolerance.
XX Multivalent vaccines have also been tried but none have successfully
XX overcome antigenic variability. The provision of further, complete
XX sequences may provide an opportunity to identify secreted or surface
XX exposed proteins that may be presumed targets for the immune system and
XX which are not antigenically variable or at least more conserved than
XX other more variable regions.
XX
SQ Sequence 92934 BP; 24051 A; 24249 C; 21902 G; 22730 T; 2 other;
```

Alignment Scores:

Pred. No.:	7.22e-06	Length:	92934
Score:	216.50	Matches:	136
Percent Similarity:	34.87%	Conservative:	76
Best Local Similarity:	22.37%	Mismatches:	208
Query Match:	4.58%	Indels:	188
DB:	21	Gaps:	21

US-09-914-168-2 (1-919) x AAA81473 (1-92934)

Qy	462	ProValIleGluThrValGluLeuThrAspGlyLeuMetAspIleSerProIleGlu	481
Db	38391	CCCGACACCCGAATCAGTCAAAATTAACCCCAATTCCTCCCTCAGACGACGAC	38450
Qy	482	PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHis	501
Db	38451	AGTGAATCAAGATATGTCGGAAGACACCTCGCGCTCATCAGCAGCAGCAGAA	38510
Qy	502	LeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArg	521
Db	38511	GTATTGGACAGGAACAGACGGGCTTCTCCCGCAAGACGCGGACACGTTAAACG	38570
Qy	522	SerIle-----LeuGlyArgIleSerAspAlaValSerAla-----	533
Db	38571	ATGCTCGCAGCAAGGCTATTTACAGCAGCAAGTCAAGCTGACGCGAAACACGCGAGCT	38630
Qy	534	-----ValAlaArgAlaIle	538

Db 38631 TATACGGTACACATCACACCGGGCCCGCCACCAAAATCGCCAACTCGCGCTCGCCATC 38690
Qy 539 LeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsn 558
Db 38691 CTCGGCAGCATCTTTTCAGACGGCAACTCGCCGAATACTACCGACCGCTGGAAAC 38750
Qy 559 ArgLysThrProAla-----AspValTyrGlnSerLysLys----- 570
Db 38751 TGGCAGCAGCGCGTAGGAGCGGATTTTCGATCAGGACAGTTGGGAAACAGCAAACTTCC 38810
Qy 570 ----- 570
Db 38811 GTCCTCGCGCGGTAAACCGCAAGCCTTACCGCTTGCCAAAGTCGCGCAATACGACGCG 38870
Qy 571 -----ValProLeuTyrValPheValAlaIleAspLysPro 582
Db 38871 GCCGTCACCCCGATACCGCCACCGCGATTGAACTGCTGCGTGACAGCGCGCGCC 38930
Qy 583 ArgAspGlyGlnIleGlyLeuGlyTyrPheGlySerAspThrGlyThrArg----- 598
Db 38931 -----ATCGCTTCGGCGACTTTGAAATCACCGGCACACAGCGTTACCCCGAA 38978
Qy 599 -----LeuVal 600
Db 38979 CAATATGCTCCGCGCTTCGCGCTTCAGCCCGGTATCGCGTACGACCTCGACCTCGTGT 39038
Qy 601 ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeu--- 619
Db 39039 CTCGACTTCCAAACAGCGCTC---GAACAACAGCGCATTTATTCGGCGCGCTCGGTACAA 39095
Qy 620 -----ArgLeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysPro 635
Db 39096 GCCGACTTCGACCGCTCAAGGCGACCGCTCCCGCTCAAGTCAGCGTAACCGAGTGC 39155
Qy 636 LeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPhe 655
Db 39156 AAACGCCCAAACTCGAAACCGGCATCCCGCTCGATTCGGGAATC----- 39200
Qy 656 GlyHisSerThrAsnGly-----PheAspLeu-SerThrArgThrLe 669
Db 39201 -----GGTTTGGCGGCAAAATCGCTTACGACTATTACAACTCTTCAACAAAGGCTAT 39254
Qy 669 uGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrPheAsnArgThrSerLe 689
Db 39255 A-----TCGGTTCGGTCTGCTGGGATATCGACAAATACG----- 39288
Qy 689 uArgTyrArgLeuAspLysLeuLys----- 697
Db 39289 -AAACACGCTTGGCGCGGCATCAGCCAGCGGCACTATCGGGGCACTACTGGACA 39347
Qy 698 -----ThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPh 712
Db 39348 AGCAAGGTTTCTTACAAACCGTTGACACCAACCAAACTCGAAACCGCGCTTCTCCGCG 39407
Qy 712 eValAsnGly-----LysProSe 718
Db 39408 GCGCTCTGGTATGTGCGGACCGCGCGGCATCGATGCCAGGCTGGGGCGGGAATTTCTC 39467
Qy 718 rGlnGluAla--LeuLeuAlaGlyValAlaVal-----HisLysThr--- 731
Db 39468 GCAGAAGCCCGAAATATCCCGGCTCGCTGCTGATTTGGGCAACAGCAGCCACGATG 39527
Qy 732 -----ValAlaAspAsnLeuValAsnProMetArgGlyTyr 743
Db 39528 CTGACCGCTCTTGGAAACGCCAGCTGCTCAACAACGCTGCTCATCCCGAAACGGCCAT 39587
Qy 744 -----ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuVal 757
Db 39588 TACCTCGACGCAAAATCGGTACGACTTTGGGCAATTTCTGCTCTCCACCGCGCTGATC 39647
Qy 758 SerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp 777
Db 39648 CGCACCTCT-----GCCGTCAGGT----- 39668

Qy 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIle-----Gln 793
Db 39669 -----TATTCTTCACGCCCGCAAAACAAAATCGCGACGTTTCATCATCGCGGACAA 39722
Qy 794 AlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAla 813
Db 39723 CGCGGTTTACACCGTTGCGCGGCAACATCCGACGCTTCTTCAGGCTGATGTTCGCGACG 39782
Qy 814 GlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLys 833
Db 39783 GCGCGCGGCTTTCGTCGCGGTTACGAACTCGACACATCGGACTTGC CGCGCCGCAAC 39842
Qy 834 GlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPhe 853
Db 39843 GGATCGGTCCTGCCGAACGCGCTCTGCTGGGACGCTGGAATACCACTGCGGCTTT 39902
Qy 854 MetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPhe 873
Db 39903 ACGCGACCTTTCGCGCGGCTGTTCCACGATATGGCGGATGCGCGCCCAATTTCAA 39962
Qy 874 ThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTyrAlaSerProValGlyGln 893
Db 39963 CGTATGAAGCTGAACACACGTTTCGGGACTGGCGTGCCTGTTTCAGCCGCTTTCGCGCG 40022
Qy 894 ValArgValAspValAlaThrGlyValLysGluGlyAsnProIleLysLeuHisPhe 913
Db 40023 TTTTCTTCGACATCGCTACGGGCACAGCGATAAG-----AAATCCGCTGGCACATC 40076
Qy 914 PheIleGlyThrProPhe 919
Db 40077 AGCTTGGGAACGCGCTTC 40094
RESULT 30
AAF21613
ID AAF21613 standard; DNA; 172325 BP.
XX
AC AAF21613;
XX
DT 13-MAR-2001 (first entry)
XX
DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:114.
XX
KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
KW ds.
XX
OS Neisseria meningitidis.
XX
PN WO200066791-Al.
XX
PD 09-NOV-2000.
XX
PF 08-MAR-2000; 2000WO-US05928.
XX
PR 30-APR-1999; 99US-0132068.
PR 08-OCT-1999; 99WO-US23573.
PR 28-FEB-2000; 2000GB-0004695.
XX
PA (CHIR) CHIRON CORP.
PA (GENO-) INST GENOMIC RES.
XX
PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Maignani V;
PI Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V, Rappuoli R;
PI Frazer CM, Grandi G;
XX
DR WPI; 2000-647603/62.
XX
PT Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections -
XX
PS Claim 7; Appendix A; 692pp; English.
XX

Db	142598	-----ATCGCCTTCGGCGACTTTGAATACACGGGCACACAGCGTTACCCCGAA	142641
Qy	599	-----LeuVal	600
Db	142646	CAAAATCGCTCTCCGGCGCTTCGGCGGTATCCGCTACGACCTCGACCTGCCTG	142705
Qy	601	ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeu---	619
Db	142706	CTCGACTTCCAAACAGGGCGCTC---GAACAACAAACGGGCATATTATCCCGCGCGTCCGTACAA	142762
Qy	620	-----ArgLeuSerGluAspLysGlyLysLysLeuIleLysLeuTyrAlaThrLysPro	635
Db	142763	GCCGACTTCGACCGCTCCCAAGCGCAGCGCGTCCCGCTCAAGCTCAGCGTACCGAGAGCTC	142822
Qy	636	LeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPhe	655
Db	142823	AAAGCCACAACTCGAAACCGGCATCCGCGCTCGATTTCGGAATAC	142867
Qy	656	GlyHisSerThrAsnGly-----PheAspLeu-SerThrArgThrLeu	669
Db	142868	-----GGTTTGGCGCGCAAAATCGCGCTATTTACAAACCTCTTCAACAAAGGCTAT	142921

[illegible]

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Qy 689 uArgTyrArgLeuAspLysLeuLys----- 697
 ::: |||||
 Db 142956 -AAACCACGGTTGCGCGCGCATACGACGCGCACTATCGGGCAACTACTGGACA 143014
 Qy 698 -----ThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPh 712
 ||||| ||||| ||||| ||||| |||||

STOCK 41 25

Db 143075 GGCCTCTGGTATGTGCCGACCGCCGGCGCATCGATCCAGCCTGGGGCGCGAATTCTTC 143134
:::|||||
Qy 718 rGlnGluAla--LeuLeuAlaGlyValAlaval-----HisLysThr--- 731
|||||||:::|||||
Db 143135 GCAGAAGCCCGAAATCCCGCGCTCGGTCTCGATTGGGCAACAGCCACGCCAGCATG 143194
Qy 732 -----ValAlaAspAsnLeuValAsnPProMetArgGlyTyrr 743
:::|||||||:::|||||||
Db 143195 CTGACCOCCTTTGGAAACGCCAGCTGCTCAACAGCTGCTCATCCGAAGCCGCAT 143254
Qy 744 -----ArgGlnArgTyrrSerLeuGluValGlySerSerGlyLeuVal 757
:::|||||||:::|||||||
Db 143255 TACTCGACGCAAAAATCGGTACGACTTTGGCACATTCCTGTCTCCACGCGCTGATC 143314
Qy 758 SerAspAlaAsnMetAlaIleAlaArgAlaGlyIleserGlyValtyrSerPheGlyAsp 777
:::|||||||
Db 143315 CGCACTCT-----GCCGTGCAAGT----- 143355
Qy 778 AsnAlatyrrGlySerAsnArgAlaHisGlnMetThrrGlyGlyIle-----Gln 793
|||||:::|||||
Db 143336 ----TATTTCCTCACGCCCGCAAAAACTCGCACGTTTCATCATCCGCGACAA 143389
Qy 794 AlaGlyTyrrIleTrpSerAspAsnPheAsnHisValProtyrArgLeuArgPheAla 813
|||||||:::|||||
Db 143390 GCGGGTTACACCGTTGCCCGCAATGCCGACGCTCTTCAGGGTGATGTTCCGACG 143449
Qy 814 GlycylaspGlnSerIleArgGlyTyrrAlaHisaspSerLeuSerProIleSerAspLys 833
|||||:::|||||||
Db 143450 GCGCGCGCGCTTCGCGCGCGGTTCGAACTCGACAGCATCGAGCTGCGCGCGCGAAC 143509
Qy 834 GlyTyrrLeuthrGlyGlyGlnValLeuAlaValcyIThrAlaGluTyrrAsnTyrrGluPhe 853
|||||:::|||||
Db 143510 GGATCGGTCTGCGCGCAACGCGCCCTCCTGGTGGCGAGCTCGGAATACCAACTGCCGTTT 143569
Qy 854 MetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrrAspLysGlyPhe 873
:::|||||
Db 143570 ACGCGCACCTTTCCGCGCGGTGTTCACGATATGGCGCATCGCGCGCAATTTCAA 143629

QY 874 ThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyGln 893
 DB 143630 CGTATGAAGCTGAACACACGCTTCGGGACCTGGGCGTGCCTTCAGCCCGCTTGCGCG 143689
 QY 894 ValArgValAspValAlaThrGlyValLysGluGluGlyAsnProIleLysLeuHisPhe 913
 DB 143690 TTTTCCTTCGACATCCCTACGGGACAGCGATAG-----AAATCCCTGGCACATC 143743
 QY 914 PheIleGlyThrProPhe 919
 DB 143744 AGCTTGGAAACGGCTTC 143761
 RESULT 31
 ABO72980
 ID ABO72980 standard; DNA: 2388 BP.
 XX
 AC ABO72980;
 XX
 DT 20-SEP-2002 (first entry)
 XX
 DE Methylococcus capsulatus surface protein-antigen D15 DNA SEQ ID NO:5.
 XX
 KW Methylococcus capsulatus; methanotropic bacterium; MopC; MopD; MopE;
 KW MopF; D15; surface protein antigen; helper protein; vaccine;
 KW epitope mapping; gene; ds.
 XX
 OS Methylococcus capsulatus.
 XX
 WO200255549-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 14-JAN-2002; 2002WO-NO00018.
 XX
 PR 12-JAN-2001; 2001NO-0000238.
 XX
 PA (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.
 PA (TIGR-) TIGR.
 XX
 PI Birkeland NK, Eidhammer I, Jonassen I, Jensen HB, Lien T;
 PI Lillehaug JR, Lossius I, Eiseen JA, Fraser CM, Durkin AS;
 PI Salzberg SL;
 XX
 WP1; 2002-557813/59.
 XX
 PT New nucleotide molecule encoding a surface exposed protein and is
 PT incorporated and expressed in Methylococcus capsulatus, useful in
 PT manufacturing vaccines that can be administered orally in animals, fish
 PT or humans
 XX
 PS Claim 1; Page 34-35; 39pp; English.
 XX
 CC The present invention describes a nucleotide molecule (I) comprising a
 CC sequence that codes for any of the 14 Methylococcus capsulatus
 CC nucleotide sequences given in ABO72976 to ABO72989. Also described is
 CC a method for producing a desired protein in a bacterial host cell,
 CC comprising: (a) transforming a bacterial host cell with a recombinant
 CC vector comprising a first nucleotide sequence selected from the
 CC group of 14 Methylococcus capsulatus, and comprising a nucleotide
 CC encoding the desired protein linked in frame to the first nucleotide
 CC sequence; and (b) culturing the transformed host cell in a medium under
 CC conditions that allow the expression of the protein. (I) can be used
 CC in vaccine production. The nucleic acids or the proteins from the
 CC present invention can be used in manufacturing vaccines that can be
 CC administered in animals, fish or humans. The method is useful in
 CC producing fusion proteins carrying foreign peptide sequences that can
 CC also be useful in epitope mapping, screening of antibody libraries and
 CC as immunogens. The present sequence represents the specifically claimed
 CC Methylococcus capsulatus surface protein-antigen D15 nucleotide sequence
 CC from the present invention.
 XX
 SQ Sequence 2388 BP: 535 A: 653 C: 686 G: 514 T: 0 other:

```
Db 799 TACGTCAATTC-----GAAATCGAGTCCACCCAGGTATCGATCAGCCCAAC 846
Qy 446 SerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGlu 465
Db 847 AAG-----AAAGAAATATATATC 864
Qy 466 ThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSer 485
Db 865 ACCATCAACGTTAAGGAGGCGCAGTCTTCAAGGTGGAGCAGGTACGGCTACCGGCAAG 924
Qy 486 AsnLeuIle-----GlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyr 503
Db 925 ACCATCGTCCACCCGACAACTCGTCCGCTCGTGGTATCGGCGCGGAGATATTTT 984
Qy 504 AspMetProAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle 523
Db 985 -----TCCAGGAAATCGCGAGCAAAACCCAGAGGCG----- 1017
Qy 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543
Db 1018 -----ATTTCGATCGTTGGGC----- 1035
Qy 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
Db 1036 -----GAGAGGGCTATATCTTCGCCCAACGTCAATATGTTGCGG 1074
Qy 564 AspValTyrGlnSerLysLysVal-----Pro 572
Db 1075 GACATCAATCAGGAGAGAACCGGTAAACATCACCTTTTCGTCGATCCGGCAACAG 1134
Qy 573 LeuTyrVal-----PheValAlaSerAspLysProArgAsp----- 584
Db 1135 GTCTACGTGAGGCGAATCAATTTCCAGGGCAATACCAAGACCGGTGAGAAAGTGTGCGC 1194
Qy 585 -----GlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyTyrArgLeuValThr 601
Db 1195 AGGGAATTCGGCAGATGAGGCGCGCTGGCGTCG-----ACCGCG 1236
Qy 602 LysPheGluHisAsn-----LeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeu 619
Db 1237 AAGATCGAGCGTTCCAAAGACCAGCGCTCGAGGCTCTCGGCTATTTCCAGGATGTCAACGTG 1296
Qy 620 ArgLeu-----SerGluAspLysLysGlyValLysLeuTyrAlaThrLys 634
Db 1297 GAGAGCGCGCGGTGCGGGAAACGAGGTACGATCGATCGACGTCAACTACAGTGTACCGAG 1356
Qy 635 ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr----- 650
Db 1357 AAGTCTTCC-----GGCAACTGACGCGCGGTGTAGGTATTTCCCACTCCAGTCCAG 1404
Qy 651 -----GlnGlnGluValPheGlyHisSer----- 658
Db 1405 GGCATCATATTCAATGCGCGGTCAACCAGGACAACATTTTCGGCAGCGGCAAGCGTGC 1464
Qy 658 ----- 658
Db 1465 AGTTCAATTTCAACACAGCCAGCATCAACACCATCTATGCTTGGGTATTTCAATCCT 1524
Qy 659 -----ThrAsnGlyPheAspLeuSerThrArgThrLeuGluHis 671
Db 1525 TACGCCACCCCTGGATGGCATCAGCAGTGGTTTCGACATCAGCTATCGCGACACCAACACC 1584
Qy 672 GluIleSerArg-----SerIleIleGlnAsnGlyGly---Trp 683
Db 1585 GGCATTTCGAACTATGTTGCCAACTACATTACCAATGTGTTCCAGGTGGCGGCAACTGG 1644
Qy 684 AsnArgThrTyr-----SerLeuArgTyrArgLeuAsp-----LysLeuLys 697
Db 1645 GGTGTGCCATCGGAGAGTTCGACAGTATTCGGACCAACCTGGACTACAGTAACACCAAG 1704
Qy 698 -ThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysPr 717
Db 1705 CTCAAGACGACCTCCC-----AGTCTTCGATCAATCAGAGCGTTTCATTGCG 1752
```

```
Qy 717 oSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuVa 737
Db 1753 GACCATGCGACGAGGTATTCGACCTATTTCCTCCAC-----TGGGT 1794
Qy 737 lAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuVa 757
Db 1795 TGGACCCATACACCCCTGAATCGGCGCATATTCG----- 1828
Qy 757 lSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPhe----- 775
Db 1829 -----CGACGAGCGGTGGCGCGCAGCGA--TTGACGGGACTGTTTGGCTCCCGTT 1877
Qy 775 ----- 775
Db 1878 CAGCACCTCGCATATTACAAGGCCAACGTGCGCTGGAGCAGTATTTTCGCTGACCCA 1937
Qy 776 -----GlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMe 788
Db 1938 GGATCTGACGTGTGTTGAATGGCGATTTCGGTTACGGC----- 1977
Qy 788 tThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrAr 808
Db 1978 ---GGCGCTATGCGAGTGGC-----GGCAACACGCGCTTCCGTTCTCG 2018
Qy 808 gLeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSe 828
Db 2019 GGAACACTTTTACGCGCGCGTCCGAATTCGGTGGGTGATACCAAGCCCAATTCACCTGG 2078
Qy 828 rProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGl 848
Db 2079 GCCAGGACAGCGCTGGCTACGCTTCGCGCGGCAACAGCAAACTGACCGGCTCGGTGGA 2138
Qy 848 uTyrAsnTyr-----GluPheMetLysAspLeuArgLeuAlaValPh 862
Db 2139 GTTGTGTTTCCGTCGCGTTTCGGGCGAGAACTGAAGAGTGTCCGCTTGGGTACCTT 2198
Qy 862 eGlyAspIleGlyAsnAlaTyr---AspLysGlyPheThrAsnAspThrLysIleGlyAl 881
Db 2199 CGTCGACGCGCGGAATGTGTTTCGTCATTCGCCCCAGCTTTCGGACCTGCGCTTCGAC 2258
Qy 881 aGlyValGlyValArgTyrAlaSerProValGlyGlnValArgValAlaThrGl 901
Db 2259 CGGATTTCCGCCAAGTGGCTATCACCTTCGGGGCCCTGATGTCTCGATCGCCCGCC 2318
Qy 901 yValLysGluGluGlyAsnProIleLysLeuHisPhe 913
Db 2319 GCTCAACTCCAGAGCGGGGAGCGGAATTCAGCATTTT 2355
RESULT 32
ABQ90024
ID ABQ90024 standard; DNA; 2391 BP.
XX
AC ABQ90024;
XX
DT 01-OCT-2002 (first entry)
XX
DE M. capsulatus gene #9 for DNA array.
XX
KW Micro array; gene; ds; differential expression; gene expression.
XX
OS Methylococcus capsulatus.
XX
PN WO20025655-A2.
XX
PD 18-JUL-2002.
XX
PF 14-JAN-2002; 2002WO-NO00019.
XX
PR 12-JAN-2001; 2001NO-0000235.
XX
PR 12-JAN-2001; 2001NO-0000239.
XX
PA (UNIF-) UNIFOB STIFTELSEN UNIV BERGEN.
```


Db 91 TTGACTTCTCCAAAAGAGCCTCTCAAGAAATCTCAAAA-----AATGAAGCTCCA 141
Qy 44 AlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrPro 63
Db 142 AAAATGAAGTTCAAAGAAATGAAGCTCAAAAAGAAACCCCAATCAATCAACGGCT 201
Qy 64 GluGlnIleGlnAlaArg-----LeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGln 82
Db 202 AAAGAAATGAAGTCAAGCTCATCTTCTTATGTCGGGCTTCTTACATCTCTGACATGCTC 261
Qy 83 AlaLeuAspValValAspPheAspGlnSerProIleSerArgIleGlyGlnSer 102
Db 262 GCTAATGAAGTGTAAAGATT-----CGTGGGCGCATATTGTG 300
Qy 103 ProProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluGlu 122
Db 301 GATTCTAAAAAATAGACACCGCTGTTTG-----CCT 333
Qy 123 LeuPheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyr 142
Db 334 TTGTTCAATCAA----- 345
Qy 143 GlnGlyGluGlnProAsnSerGluValValProProThrLeuGluProGluLysPro 162
Db 345 ----- 345
Qy 163 GlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArg 182
Db 346 GGGTATTTAAAGACCTTTATGCC---ACTTTGAGCGCGCAT----- 387
Qy 183 LeuLysAlaLysPheTyrGlnSerSerGlnSerGly-----GluThrSerAlaIleGly 200
Db 388 TTAGAGTTTCATTTTGATGAAAAAGCCAGGATTGCGGGGTAGAAATCAAGGGTTATGGG 447
Qy 201 SerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThr 220
Db 448 ACTGAAAGGAAAAAGACGGCTTAAATCCCAATGGGGATCAAAAAGGGCAGACCTTT 507
Qy 221 GlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeu--- 239
Db 508 GATGAGCAAAAATTAGAG-----CATGCTAAAACGGCTTAAAA 546
Qy 240 ValAlaAlaArgAlaValAlaGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIle 259
Db 547 ACCGCTTTAGAGGGCGAGGGCTATTAT----- 573
Qy 260 GlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAla 279
Db 574 GGGACCGTG----- 582
Qy 280 ValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAsp----- 297
Db 583 GTGGAGGTGCGCACAGAAAGTCAGTGAGGTGCATTATTGATCGTGTTCATGTGAAT 642
Qy 298 ---GluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLys 316
Db 643 AGGGGGGTAGCATTTATATCAACAATCCATTTATAGGGAAGCGCAAAATTAAGACGC 702
Qy 317 AsnLeuIleGluAsnAlaSerAlaGluHis-----Gly 327
Db 703 CGCATTTGTAATCTTTGAGTGCAGAACAGCAACGAGATTTCATGGGTGGATGTGGGCG 762
Qy 328 TyrPheAspGlyArg----- 332
Db 763 TTGAATGACGGAAATTCGGTTTAGATCACTAGATAATACGATTCTATCGGTATCCAAAGT 822
Qy 333 -----TriLeuAspArgSerValasp-----ValIleLeuProAsp 344
Db 823 GTGTATATCGGTAGGGTTTACTTAGACGCTCATATTTCTTCGCCCTTTTGTGAAACGGAT 882
Qy 345 ---AsnThrAlaAspValSerLeuIleTyrAspThr-----GlyThrGlnTyrArgPhe 361
Db 883 TTTTACCCATGACGCTTAAGCTTCATTATATAAGTCAAAAGGAGGGATCCAAATACAGGATT 942

Qy 362 AspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAsp 381
Db 943 TCAGACATTCTTA-----ATAGACATTGACAAACCGG 972
Qy 382 LysLeuProValLysArgGluLeuGluGlnLeuLeuThrValAsnMetGlyGluAla 401
Db 973 GTAGTCCCTTAAAA-----ACCTTAGAAAAGCGCTTAAAGTAAAAAGGAAGATGTC 1026
Qy 402 TyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeu-----IleAlaThrArg 418
Db 1027 TTTAATATTGACATTTAAGAGCGGATCGCAAAATTTTAAAAACCAATCGCCGATAAG 1086
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 1087 GGTATGCGTTTGGC-----GTGCTGAAGCCAGACTGGATAAAA----- 1125
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSer 458
Db 1126 -----CATGAAAAA 1134
Qy 459 Thr---LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAsp--- 476
Db 1135 AACGGGCTTGCAAGTCATTTATCGTATTGAAGTGGCGGATATGGTGTATATCAATGAT 1194
Qy 477 -----IleSerProIleGluPheSerAlaSerAsnLeuIle----- 488
Db 1195 GTCATCATTTTCAGGGAACCGCACGCGATAGGATCATTAGAAGGAGTATTATTGTA 1254
Qy 489 -----GlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetPro 506
Db 1255 GGGCTAAGGATAAACAACCTTGACCAAACTGAGA----- 1290
Qy 507 AspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArg 526
Db 1291 -----AATCCGAAAATCTTTAAGCGCT-----TTAGGATTC 1323
Qy 527 IleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu 546
Db 1324 TTCCTAAAGTCAAAATTTGAAGAAAAAGGCTT-----AATAGCTCA 1365
Qy 547 ValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr 566
Db 1366 CTCATGGATTTA----- 1377
Qy 567 GlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGln 586
Db 1378 -----TTAGTGCCGTAGAAAGCGGCGTACTGGGCAG 1410
Qy 587 Ile-----GlyLeuGlyTyrpGlySerAspThrGlyThrArgLeuValThrLysPheGlu 604
Db 1411 TTGCAATTTGGGTAGGCTATGGCTCTTATGGAGGCTTATGCTT-----AAT 1458
Qy 605 HisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAsp 624
Db 1459 GGGAGCTTGACGCAAGAAACCTTTTGGCACAGGCAAGCATGAGCTTGTATGCTAAC 1518
Qy 625 -----LysLysGlyValLysLeuTyrAlaThrLysPro----- 635
Db 1519 ATCGCTACAGGGGGGTAGATCTTATCCGGGCATGCCAAAAGGCGGGCGTATGTTT 1578
Qy 636 -----LeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
Db 1579 GCCGGGAATTTGAGCTTGACTAATCCA----- 1605
Qy 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeu--- 669
Db 1606 -----AGGATTTT-----GACAGCTGGTATAGCTCTACGATCAACCTTTAT 1647
Qy 670 ---GluHisGluIleSerArgSerIleIleGlnAsnGlyGly----- 682
Db 1648 GCGGATTACAGATAAGACTACCAATACATCCAAAGCGGGGGCTTTGGGTGAATGTC 1707


```

QY 683 -----TrpAsnArgThrTyr---SerLeuArgTyrArgLeuAspLysLeuLys 697
Db 1708 GGCGCGATGCTGGGTAAAGACCCATGTGAGCTTACCTGAATGAATGACAAA 1767

QY 698 ThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheValasnGly----- 715
Db 1768 CTCTTGGTTTCAGCAGCCCTTTATACACCGGTACTATTCTCTCTGTATGAAGTGGTT 1827

QY 716 LysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAla----- 733
Db 1828 TCTCCAAGCAATGTTCTACCCCGCATCGTGTATTATCAATCGCTATCAGCGGTAAA 1887

QY 734 -----AspAsnLeuValasnPro-----Met 740
Db 1888 ACCCCCTTACAACCTGAAGAGCTGTTCTAGTCTCTGGAGCATCACCACTTCCACGAGAATA 1947

QY 741 ArgGlyTyrArgGlnArg-----TyrSerLeuGluVal 751
Db 1948 AGAGGTATTGGGATAGGGATTACCATACCCCTATCACCAGCTCTTTCACCCCTGATGTG 2007

QY 752 Gly-----SerSerGlyLeuValSerAspAlaAsn 761
Db 2008 AGCTATGACAAACCGATGATTATTACTCCCTAGAAATGGGCTTATCTTTAGT----- 2061

QY 762 MetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp----- 777
Db 2062 -----TCCATGCGAGATGTTGGCTTGCACAGCTCTGCACGCTCAATTTCTTG 2112

QY 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGly----- 791
Db 2113 ACGGGTTAGCGGGAATGTCCCTACACCAAAAGTTATGGTAAATTCGCCGCTTACCAC 2172

QY 792 -----IleGlnAlaGlyTyr 796
Db 2173 CATTTGCAAAATATTATTATGATAGATTGATCGCTGCTTTAAACGCAAGCAGGTAT 2232

QY 797 IleTrpSerAspAsnPhe---AsnHisValProTyrArgLeuArgPhePheAlaGlyGly 815
Db 2233 ATCTTAGGTATACACCGCATGATTACTTGCCTTAAACTCCACCTCTACATGGGGGGC 2292

QY 816 AspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyr 835
Db 2293 GTAAACACCGTGAGCGCTTTAGAACGGATCGTTACTCTCTAAAGATGATTTGCGTTG 2352

QY 836 LeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLys 855
Db 2353 TGGCTTGGAGCGCATGGGATTTTACCCTCTCTACTGAATTTGAGCTAGGCTGCTAAAG 2412

QY 856 -----AspLeuArgLeuAlaValPheGlyAspIleGly----- 866
Db 2413 GCGGCTAAATGCGCTAGCGTGGTGTGTTTGTACTTTGGTTTCTTAACCTTTAAACCCCA 2472

QY 867 -----Asn 867
Db 2473 ACTAGAGGAGTGTTCCTATACGCTCTGTTACGACAGCGAATTTTAAAGATTATGGC 2532

QY 868 AlaTyrAspLysGlyPheThrAsnAspThr---LysIleGlyAlaGlyValGlyValArg 886
Db 2533 GTTATAGGGCTGGTGTGAAGACGCACTTGGAGGCTTCCACAGGCTTGCAGATTGAA 2592

QY 887 TrpAlaSerProValGly 892
Db 2593 TGGATTTCGCCCATGGGG 2610

RESULT 34
AAV52012
ID AAV52012 standard; DNA; 3012 BP.
XX
AC AAV52012;
XX
DT 09-NOV-1998 (first entry)
XX
DE Helicobacter polypeptide GHPO 107 DNA.

```

```

XX GHPO 107; infection; therapy; diagnosis; vaccine; gastritis; ulcer;
KW ss.
XX Helicobacter pylori.
XX OS
XX Key Location/Qualifiers
FH CDS 142..2685
FT /*tag= a
XX
XX PN W09821225-A1.
XX PD
XX PF 22-MAY-1998.
XX PF 14-NOV-1997; 97WO-US21353.
XX PR 29-JUL-1997; 97US-0902615.
XX PR 14-NOV-1996; 96US-0749051.
XX PR 01-APR-1997; 97US-0831309.
XX PR 01-APR-1997; 97US-0833457.
XX PR 01-APR-1997; 97US-0834705.
XX PR 24-JUN-1997; 97US-0881227.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
XX (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX
XX PI Al-Garawi A, Haas R, Kleanthous H, Meyer T, Miller C;
XX PI Odenbreit S, Tomb J;
XX
XX WPI; 1998-297855/26.
XX P-PSDB; AAW71477.
XX
XX Helicobacter polynucleotide and polypeptide sequences - useful to
treat or prevent gastrointestinal infection
XX
XX Disclosure; Page 82-86; 362pp; English.
XX
XX This polynucleotide encodes Helicobacter pylori polypeptide
GHPO 107 (see AAW71477). 85 Polynucleotides (see AAV52009-V52093),
encoding claimed Helicobacter polypeptides (see AAW71474-W7158) have
been identified in the H. pylori genome using a transposon shuttle
mutagenesis method in which TnMax9 mini-blam was used for
insertional mutagenesis of an H. pylori gene library established in
Escherichia coli. The invention provides: methods for producing
these polypeptides in recombinant host systems, and related
expression cassettes, vectors and transformed or transfected host
cells; live vaccine vectors that contain polynucleotides of the
invention and which can be used to prevent or treat Helicobacter
infection; therapeutic and/or prophylactic methods involving
administration of polynucleotide molecules (either in naked form
or delivered with a delivery vehicle), polypeptides or monospecific
antibodies; methods for detecting the presence of Helicobacter
in samples using e.g. polynucleotide molecules; and methods for
purifying polypeptides of the invention.
XX
SQ Sequence 3012 BP; 896 A; 555 C; 735 G; 826 T; 0 other;

```

```

Alignment Scores:
Pred. No.: 6,39e-07 Length: 3012
Score: 205.50 Matches: 201
Percent Similarity: 32.89% Conservative: 143
Best Local Similarity: 19.22% Mismatches: 319
Query Match: 4.35% Indels: 383
DB: 19 Gaps: 51

```

US-09-914-168-2 (1-919) x AAV52012 (1-3012)

```

QY 24 MetThrSerGlnAlaLeuAlaGlnGlnAsnProAlaAsnIleAsnHisValPro 43
XX ::::::::::: ||| ||| ::|
XX
Db 25 TTGACTTCCAAAGAGCGCTCTCAAGAATCTCAAAAA-----AATGAAGTCCA 75
XX
QY 44 AlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyAsnProValLeuLeuThrPro 63
XX

```

Db 76 AAAAAAGAGTTCAACTCAAAAAGAAACCCCAAAACCCCAATCAATCAAAACGCCT 135
Qy 64 GluGlnIleGlnAlaArg- --LeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGln 82
Db 136 AAAGAAATGAAGACCAAGTCCATCTTATGTCGGGCTTTCTTACATCTCTGACATGCTC 195
Qy 83 AlaLeuAspValValAsnPheAspAspGlnSerProIleSerArgIleGlyGlnGlnSer 102
Db 196 GCTAATGAATTTGTAAGATT-----CGTGTGGCGCATATTCTG 234
Qy 103 ProProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluGlu 122
Db 235 GATTCTAAAAAATAGACACCGCTGTTTG-----GCT 267
Qy 123 LeuPheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyr 142
Db 268 TTGTTCATCAAA----- 279
Qy 143 GlnGlyGluGlnProAsnSerGluValValProProThrLeuGluProGluLysPro 162
Db 279 ----- 279
Qy 163 GlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAlaLysValProArg 182
Db 280 GGGTATTTTAAAGACGTTTATGCC---ACTTTTGAAGCGGCATA----- 321
Qy 183 LeuLysAlaLysPheTyrGlnSerSerGlnSerGly-----GluThrSerAlaIleGly 200
Db 322 TTAGATTTCATTTTGATCAAAAAGCCAGGATTGCGGGGTAGAAATCAAGGGTTATGGG 381
Qy 201 SerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThr 220
Db 382 ACTGAAAAAGAAAAGCGGCTTAAATATCCCAATGGGGATCAAAAAGGGCCACACCTTT 441
Qy 221 GlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeu--- 239
Db 442 GATGAGCAAAATTTAGAG-----CATGCTAAAAACGGCTTTAAAA 480
Qy 240 ValAlaAlaArgAlaValAlaGlyTyrTyrAspIleAspLeuSerIleArgAsnSerIle 259
Db 481 ACCGCTTTAGAGGGCGCGGCTATTAT----- 507
Qy 260 GlyGluValAlaValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAla 279
Db 508 GGGAGCGTC----- 516
Qy 280 ValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAsp----- 297
Db 517 GTGGAGGTGGCGACAGAAAGGTCAAGTGGGTGCATTTATTGATGCTGTTGATGTGAAT 576
Qy 298 ---GluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLys 316
Db 577 AGGGGGGATAGCATTTATATCAAAATCCATTTTATGAGGAAGCCGCAAAATTAAGCCG 636
Qy 317 AsnLeuIleGluAsnAlaSerAlaGluHis-----Gly 327
Db 637 CGCATGATTGAATCTTTGAGTCGGAACAAAGCAAGACAGAGATTTCATGGGTGGATGTGGGC 696
Qy 328 TyrPheAspGlyArg----- 332
Db 697 TTGAATGACGGAAATTCGGTTTAGATCAACTAGATAGCATTCATCGGTATCCAAAGAT 756
Qy 333 -----TrpLeuAspArgSerValAsp-----ValIleLeuProAsp 344
Db 757 GTGTATATCGGTAGGGTTACTTAGACGCTCATATTTCTCGCCCTTTTGTGAAACCGAT 816
Qy 345 ---AsnThrAlaAspValSerLeuIleTyrAspThr-----GlyThrGluTyrArgPhe 361
Db 817 TTTTCTACCCATGACGCTTAAGCTTCATTATATAAGTCAAAAGAGGGGATCCAAATACAGGAT 876
Qy 362 AspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProasp 381
Db ----- 381

Db 877 TCAGACATTTTA-----ATAGAGATTGACAACCCG 906
Qy 382 LysLeuProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMetGlyGluAla 401
Db 907 GTAGTCCCTTAAAA-----ACCTTAGAAAAGCGCTTAAAGTGAAGAGGAAGATGTC 960
Qy 402 TyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeu-----IleAlaThrArg 418
Db 961 TTTAATAATTGAGCATTTAAGAGCGGATCGCAAAATTTTAAAAACCCGAAATCCCGCATAG 1020
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 1021 GGTATATGCTTTGCG-----GTGGTGAAGCCAGACATTTGGATAAAA----- 1059
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSer 458
Db 1060 -----GATCAAAAA 1068
Qy 459 Thr---LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAsp--- 476
Db 1069 AACGGCTTGTGAAAGTCATTTATCGTATTGAAGTGGCGCATATGCTGTATATCAATGAT 1128
Qy 477 -----IleSerProIleGluPheSerAlaSerAsnLeuIle----- 488
Db 1129 GTCATCATTTTCAGGAACCGCACGCGATAGGATCATTTAGAGGGAGTTATTGTTA 1188
Qy 489 -----GlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetPro 506
Db 1189 GGGCTAAGGATATACTACACTTGACCAACTGACA----- 1224
Qy 507 AspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArg 526
Db 1225 -----AATTCGAAAAATCTTTAAAGCGT-----TTAGGATTC 1257
Qy 527 IleSerAspAlaValSerAlaValAlaAlaIleLeuProAspGluSerGluAsnGlu 546
Db 1258 TTCTCTAAAGTCAAAATTTGAAGAAAAAGGGTT-----AATAGTCA 1299
Qy 547 ValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr 566
Db 1300 CTCATGGATTTA----- 1311
Qy 567 GlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGln 586
Db 1312 -----TTAGTGCAGGTAGAAGGGCGCTACTGGCAG 1344
Qy 587 Ile-----GlyLeuGlyTrpGlySerAspThrGlyThrArgLeuValThrLysPheGlu 604
Db 1345 TTGCATTTGGGTTAGGCTATGGCTCTTATGGAGGGCTTATGCTT-----AAT 1392
Qy 605 HisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAsp 624
Db 1393 GGGAGCGTGAGCGAAGAAACCTTTTGGCAGGCAAGCATGAGCTTGTATGCTAAC 1452
Qy 625 -----LysLysGlyValLysLeuTyrAlaThrLysPro----- 635
Db 1453 ATCGCTACAGGGGGGTAGATCTTATCCGGCATCGCCAAAAGGAGCGGCGTATGTT 1512
Qy 636 -----LeuSerProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
Db 1513 GCCGGGAATTTGAGCTTGACTAATCCA----- 1539
Qy 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeu--- 669
Db 1540 -----AGGATTTT-----GACAGCTGGTATAGCTCTACGATCAACCTTTAT 1581
Qy 670 ---GluHisGluIleSerArgSerIleIleGlnAsnGlyGly----- 682
Db 1582 GCGGATTACAGATAAGCTACCAATACATCCAAACAGCGGGGGCTTTGGGTGAATGTC 1641
Qy 683 -----TrpAsnArgThrTyr---SerLeuArgTyrArgLeuAspLysLeuLys 697
Db 1642 GGGCGCATGCTGGGGTAATAGAACCCATGTGAGCTTAGGTTAGCTTAACCTTTACCAAA 1701

```
Qy 698 ThrGlnAlaProGluThrTrpGlnAspLeuProValAspPheValAsnGly----- 715
Db 1702 CTCCTTGTTTCAGCAGCCCTTTATACACCGCTACTATCTCTGTTAATGAAGTGGTT 1761
Qy 716 LysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAla----- 733
Db 1762 TCTCCAAGCAATGTTCTACCCGCGCATCGTGATTATCAATCGCTTATCAGCGGTA 1821
Qy 734 -----AspAsnLeuValAsnPro-----Met 740
Db 1822 ACCCCCTTACAACCTGAAGCTGTTCTAGTCTGGACGCGATCACCACCTTACCAGAAATA 1881
Qy 741 ArgGlyTyrArgGlnArg-----TyrSerLeuGluVal 751
Db 1882 AGAGGTATTTGGGATAGGGATTACCATACGCCATCACCAGCTCTTTTACCCCTTGATGTG 1941
Qy 752 Gly-----SerSerGlyLeuValSerAspAlaAsn 761
Db 1942 AGCTATGACACACCGCATGATTATTACTTCCTAGAAATGGGTTATCTTTAGT----- 1995
Qy 762 MetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp----- 777
Db 1996 -----TCCTATGCGAGCATGCTGGCTTGCCACAGCTCTGGCAGCGCTCAATCTTGG 2046
Qy 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGly----- 791
Db 2047 AACGGTTAGCGGGAATGCCGTAACACCAAAAGTTTATGTAATTCGCCGCTTACCAC 2106
Qy 792 -----IleGlnAlaGlyTyr 796
Db 2107 CATTTGCAAAATATTATTATGATAGATTGTCGCTCGCTTTAAACGCAAGGAGGTAT 2166
Qy 797 IleTrpSerAspAsnPhe---AsnHisValProTyrArgLeuArgPheAlaGlyGly 815
Db 2167 ATCTTTAGGTATACACCGATGATTACTTGCCTTAACCTCCACCTTCTACATGGGGGC 2226
Qy 816 AspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyr 835
Db 2227 GTAACCAACGGTGAGAGCTTTAGGACGGATCGGTTACTCTAAAGATGAGTTTGGCTTG 2286
Qy 836 LeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGlyValTyrAsnTyrGluPheMetLys 855
Db 2287 TGGCTTGGAGCGATGGGATTTTACCGCTTCTACTGAATAGCTATGGGCTCTAAAG 2346
Qy 856 -----AspLeuArgLeuAlaValPheGlyAspIleGly----- 866
Db 2347 GCGCTAAATGCGCTTAGCGTGCTTTTGGCTTTGACTTTGTTTCTTAACCTTTAAACCCCA 2406
Qy 867 -----Asn 867
Db 2407 ACTAGAGGGAGTTTCTATACGCTCTGTTACGACAGCGAATTTTAAAGATTATGGC 2466
Qy 868 AlaTyrAspLysGlyPheThrAsnAspThr---LysIleGlyAlaGlyValGlyValArg 886
Db 2467 GTTATAGGGCTGGGTTTGAAGACCCATCGGAGGCTTCCACAGCGCTTCCACAGATTGAA 2526
Qy 887 TrpAlaSerProValGly 892
Db 2527 TGGATTTCGCCCATGGGG 2544
RESULT 35
AAV25138
ID AAV25138 standard; DNA: 2778 BP.
XX
AC AAV25138;
XX
DF 13-JUL-1998 (first entry)
XX
DE H. pylori cell envelope OMP ORF 05ep10815_4719175_c1_l15.
XX
KW Cytoplasmic; vaccine; prevention; treatment; infection; envelope;
KW identification; binding compound; bacteria; life cycle; activator;
```

```
KW inhibitor; duodenal ulcer disease; chronic gastritis; diagnosis;
KW cell envelope outer membrane protein; OMP; ds.
XX
OS Helicobacter pylori.
XX
FH Key Location/Qualifiers
FT CDS 1..2778
FT /tag= a
XX
PN WO9737044-A1.
XX
PD 09-OCT-1997.
XX
PF 27-MAR-1997; 97WO-US05223.
XX
PR 06-DEC-1996; 96US-0761318.
XX
PR 29-MAR-1996; 96US-0625811.
XX
PR 02-APR-1996; 96US-0758731.
XX
PR 25-OCT-1996; 96US-0736905.
XX
PR 28-OCT-1996; 96US-0738859.
XX
PA (ASTR ) ASTRA AB.
XX
PI Alm RA, Smith D;
XX
DR WPI; 1997-503122/46.
XX
DR P-PSDB; AAW55729.
XX
PT Helicobacter pylori nucleic acid sequences and encoded
PT polypeptides(s) - useful in vaccines to treat or prevent H. pylori
PT infection and for diagnosis of H. pylori infection
XX
PS Claims 5.6.21; Page 929,930; 1145pp; English.
XX
XX This sequence encodes a H. pylori cell envelope outer membrane
XX protein having a terminal Phe residue.
XX The protein may be used in a vaccine to prevent or treat H. pylori
XX infection or to identify H. pylori life cycle activators or inhibitors. The
XX useful as potential H. pylori life cycle activators or inhibitors. The
XX DNA and probes derived from it may be used for the identification of
XX H. pylori in a sample and the diagnosis of H. pylori infection. Nucleic
XX acid sequences complementary to the DNA act as antisense sequences and
XX can be used to prevent the translation of H. pylori mRNA. Antibodies
XX against the protein can be used in immunoassays to evaluate the abundance
XX and distribution of H. pylori-specific antigens. The genomic sequence of
XX H. pylori (ATCC 55679) was determined from overlapping contigs generated
XX by mechanically shearing the bacterial DNA. The sequences were analysed
XX for ORF of at least 180 nucleotides, and the predicted coding regions
XX defined by computer evaluation. To identify likely H. pylori antigens for
XX vaccine development, the amino acid sequences predicted from various ORF
XX were analysed for significant homology to other known or exported
XX membrane proteins. Having identified and determined the sequences of
XX interest, particular regions can be isolated from H. pylori by PCR
XX amplification for recombinant polypeptide production, e.g. in E. coli
XX hosts.
SQ Sequence 2778 BP; 813 A; 514 C; 681 G; 770 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 1 61e-06 Length: 2778
Score: No.: 200.00 Matches: 206
Percent Similarity: 34.32% Conservatve: 143
Best Local Similarity: 20.26% Mismatches: 318
Query Match: 4.23% Indels: 352
DB: 18 Gaps: 51
```

US-09-914-168-2 (1-919) x AAV25138 (1-2778)

```
Qy 47 ThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGlnGlnIle 66
Db 128 ACGGCTCTAAACCAACG-----ATTTGGCTTCTCCAAAGAAACCC 169
Qy 67 GlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspVal 86
```

Db	170	CAAAGAGCGCTCAAAAAATGAAGTCTCAAAACGAACCT-----			-----	208
Qy	87	ValAsnPheAspAspGlnSerProIleSerArgIleGly-----			-----	104
Db	209	-----CTCAATCCATCAACGCCCTAAGAAATGAAAGTCAAGTCAATTT			-----	253
Qy	105	LeuGlyLeuAspMetSerValIleGluGluThrProLeuSerLeuGluLeuPhe			-----	124
Db	254	CTTAT-GTCGGGCTTTCTTAC-----ATGCTCTGCATGCTC			-----	288
Qy	125	AlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGly			-----	144
Db	289	GCTAATGAATAATCCAAGATTCGCGT-----GGC			-----	318
Qy	145	GluGlnProAsnSerGluValValProProThrLeuGluProGluLysProGlyLeu			-----	164
Db	319	GATATGGTGGATTCTAAAAAATAGACCGCTGTTTAGCTTCTTCAACCAAGGTAT			-----	378
Qy	165	IleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLys			-----	184
Db	379	TTTAAAGACGCTTTATGCC--ACTTTGAAAACGGCAAT-----TTAGAG			-----	420
Qy	185	AlaLysPheTyrGlnSerSerGlnSerGly-----GluThrSerAlaIleGlySerSer			-----	202
Db	421	TTTCATTTTGTGATAAAGCCAGGATGCGCGGGTAGAAATCAAGGGTATGGACGTAA			-----	480
Qy	203	HisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGlu			-----	222
Db	481	ANGMAAAGACGGCTAAATCCCAATGGGGATCAAAAGGGCGACACCTTTGATGAG			-----	540
Qy	223	SerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeu--ValAla			-----	241
Db	541	CAAAAATTAGAG-----CATGCTAAAACGGCTTTAAAAACGGCT			-----	579
Qy	242	AlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGlu			-----	261
Db	580	TTAGAGGGCAGGGCTATTAT-----GGGAGC			-----	606
Qy	262	ValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGlu			-----	281
Db	607	GTG-----GTGGAG			-----	615
Qy	282	ValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAsp-----Glu			-----	298
Db	616	GTGCGCACAGAAAGGTCAGTCAGGAGGCGTTATTGATCGTGTGTTGATGTCGAATAGGGG			-----	675
Qy	299	ValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeu			-----	318
Db	676	GACAGTATTATTATCAACAATCCATTATTATAGGAAGCGATAAATTAACACGGCGTGTG			-----	735
Qy	319	IleGluAsnAlaSerAlaGluHis-----GlyTyrPhe			-----	329
Db	736	ATTGAATCTTTGACGGCGAACAGCAGCGGATTTCATGGCTGGATGTCGGGGCTGAAT			-----	795
Qy	330	AspGlyArg-----			-----	332
Db	796	GACGGGAATTCGCGTTAGATCAATTAGAATACGATTCTTTGCGGTATCCAAGATGTGTAT			-----	855
Qy	333	-----TrpLeuaspArgSerValasp-----ValIleLeuProAsp---Asn			-----	345
Db	856	ATGCGTAGGGGTACTTAGACGCTCATATTCTCGCCCTTTTGTGAAACGGATTTTTC			-----	915
Qy	346	ThrAlaAspValSerLeuIleTyrAspThr-----GlyThrGlnTyrArgPheAspGlu			-----	363
Db	916	ACCATGACGCTAAGCTCCATTATAGGTCAAGAGGGGATCCATACAGATTTTCAGAT			-----	975
Qy	364	ValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeu			-----	383
Db	976	ATTITTA-----ATAGAGATTGACAACCCGGTAGTC			-----	1005
Qy	384	ProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsn			-----	403

Db	1006	CCCTTAAAA- - - - -	ACCTTAGAAAAAGCGCTTAAAGTTTAAAGAGGAAGATGCTCTTTAAT	1051
Qy	404	LeuGlnAlaValArgAlaLeuSerAsnAspLeu- - - - -	IleAlaThrArgTyrPhe	420
Db	1060	ATTGAGCATTTTAAAGCGGATCGCAAAATTTTAAAAACCGAAATCCCGATAGGCGCTAT	1119	
Qy	421	AsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnVal	440	
Db	1120	CGGTTTGGC- - - - -	GTGGTGAAGCCAGACTTGGATAAA- - - - -	1152
Qy	441	SerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThr- - -	459	
Db	1153	- - - - -	GACGAAAAACGGG	1167
Qy	460	LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro	479	
Db	1168	CTTGTGAAGTCATTATTCGTATTGAAGTGGCGAT- - - - -	ATGTGTCATATCAATGAT	1221
Qy	480	IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaLysAla	499	
Db	1222	GTCATCATTTTCAGGAAACAGCGCAGCGATAGATC- - - - -	ATTAGG	1266
Qy	500	ArgHisLeuTyrAspMetProAspArg- - - - -	ValLeuAlaIleAsnHisAsp	516
Db	1267	AGGGAAATCTTACTAGGCGCTTAAGATAATAACAACTTGACCAACTGAGAAATCCGAA	1326	
Qy	517	AspGlyValAsnArgSerIleLeuGlyArgGlyIleSerAspAlaValSerAlaValAlaArg	536	
Db	1327	AATTCCTTCAGCGGT- - - - -	TTAGSGTTTCTCTAAAGTCAAGATTGAAGAAAAAAGG	1380
Qy	537	AlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeu	556	
Db	1381	GTC- - - - -	AATAGCTCATTCATGGATTG- - - - -	1404
Qy	557	AlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPhe	576	
Db	1405	- - - - -	TTA	1407
Qy	577	ValAlaSerAspLysProArgAspGlyGlnIle- - - - -	GlyLeuGlyTrpGlySerAsp	594
Db	1408	GTGAGCGTAGAAGAGGGCGCACCGCGCAGTTCGAATTCGGGTTGGCTATGCTCTTAT	1467	
Qy	595	ThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGln	614	
Db	1468	GGAGGGCTCATGCTT- - - - -	AATGGAGCGGTGAGCGAAAGGAATCTTTTGGC	1515
Qy	615	AlaGlyAlaGluLeuArgLeuSerGluAsp- - - - -	LysLysGlyValLysLeuTyrAla	632
Db	1516	ACAGGCAAGCATGAGCTTGATGTGTACATTGCCACAGGGGGGTAGATCTTATCCG	1575	
Qy	633	ThrLysPro- - - - -	LeuSerHisProLeu	640
Db	1576	GGCATGCCAAAGGGCGCGGTATGTTGCCGGGAATTTGAGCTTGACTAATCCA- - -	1632	
Qy	641	AsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPheGlyHisSerThrAsn	660	
Db	1633	- - - - -	AGGATTTTT- - - - -	1644
Qy	661	GlyPheAspLeuSerThrArgThrLeu- - - - -	GluHisGluIleSerArgSerIleIle	678
Db	1645	AGCTGGTATAGCTCTACGATCAATCTTTATGCGGATTACAGGATACAGTACCAATACAT	1704	
Qy	679	GlnAsnGlyGly- - - - -	TrpAsnArgThrTyr- - -	687
Db	1705	CACAAAGCGCGGGCTTTGGGTGAATGTCGGGCGCATGCTGGGTAAAGAACCCATGTG	1764	
Qy	688	SerLeuArgTyrArgLeuAsp- - - - -	LysLys- LysThrGlnAlaProProGluThrTr	705
Db	1765	AGCTTAGGTATAAATCTGAATGTTACCAAACTCCTTGGTTTCAGCAGCCCTTATACAAC	1824	
Qy	705	pGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGln	725	
Db	1825	CGCTACTACTCTCTTAATGAAGTG- GCCTCTCCAAGCAATGTTTCCACACCCGCATC	1883	

Db	1381	GTC	----		----	-AATAGCTCATGTATGGATTG-	----	1404
Qy	557	AlaAsnArgLysThrProAlaAspValTyrGlnSerLysValProLeuTyrValPhe	576							
Db	1405	----						-----TTA	1407	
Qy	577	ValAlaSerAspLysProArgaspGlyGlnIle	594					GlyLeuGlyTyrTrpGlySerAsp		
Db	1408	GTAGAGCTAGAAGAGGGCGCCACCGGGAGTTGCAATTCGGTGTGGCTATGGCTTAT	1467							
Qy	595	ThrClyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgaspGlyTyrGln	614							
Db	1468	GGAGGGCTCATGCTT	1515					-----AATGGAGCGTGACGAAAGGAATCTTTTGGC		
Qy	615	AlaGlyAlaGluLeuArgLeuSerGluAsp	632					-----LysLysGlyValLysLeuTyrAla		
Db	1516	ACAGGGCAAGCATGAGCTTGTATGCTAACATTGCCACAGGGGGGTAGATCTTATCCG	1575							
Qy	633	ThrLysPro	640					-----LeuSerHisProLeu		
Db	1576	GGCATGCCAAAGAGGGCGGGGTATGTGTTCGGCGGAATTTGAGCTTGACTAATCCA	1632					-----GAC		
Qy	641	AsnAspClnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsn	660					-----AGGATTTT		
Db	1633	----						-----GAC	1644	
Qy	661	GlyPheAspLeuSerThrArgThrLeu	678					-----GluHisGluIleSerArgSerIleIle		
Db	1645	AGCTGGTATAGCTCTACGATCAATCTTTATCGGGATTCACAGATAAGCTACCAATACATC	1704							
Qy	679	GlnAsnGlyGly	687					-----TrpAsnArgThrTyr		
Db	1705	CAACAGCGGGGGCTTTGGGTGAATGTCGGCGCATCTGGGTAAATAGAACCATGTG	1764							
Qy	688	SerLeuArgTyrArgLeuAsp	705					-----LysLeu-LysThrGlnAlaProProGluThrTyr		
Db	1765	AGCTTAGGTATACCTTGATGTATACCAACATCTTGTGTTTCAGCAGCCCTTATACAC	1824							
Qy	705	pGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaG	725					-----LysLeuLeuAlaG		
Db	1825	CGCTACTATTCCTCTGTTAATGAAGTG-CCCTCTCCAGGCAATGTTCCACACCCGATC	1883					-----ThrValAlaAspAsnLeuValas	738	
Qy	725	yValAlaValHisLys						-----ThrValAlaAspAsnLeuValas		
Db	1884	GGTGATTAACAACCGCTTATCAGCGGTAGAACTCCATTGGTCTCTGAAAGCTGTTCTAG	1943							
Qy	738	nPro	746					-----MetArgGlyTyrArgGlnArg		
Db	1944	TCCTGGAGCGATCACCACTTCACCAGAAATAAAGATATTTGGGATAGGGATTACCACAC	2003							
Qy	747	----	752					-----TyrSerLeuGluValGly		
Db	2004	GCCTATCACCATGCTTTCACCTTGATGTGAGCTATGACAAACCCGATGATTAATATT	2063					-----SerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerG	771	
Qy	753	----						-----AsnAlaTyrGlySerAsnArgAlaHisG	787	
Db	2064	CCCTAGAAATGGGGTTATCTTAGT	2108					-----TCCATGGCAATGCTCGG		
Qy	771	yValTyrSerPheGlyAsp						-----AsnAlaTyrGlySerAsnArgAlaHisG		
Db	2109	TTTGCCAAAGCTCTGGCAGCGCTAATCTTGTGGAAACGGTTAGCGGGAAATGTCGTAACAC	2168							
Qy	787	nMetThrGlyGly	791					-----		
Db	2169	CAAAGTTATGTGTAATTCGCGCTTACACCATTTGCCAAATAATATTTATGATAGATT	2288					-----IleGlnAlaGlyTyrIleTrpSerAspAsnPhe	805	
Qy	792	----						-----IleGlnAlaGlyTyrIleTrpSerAspAsnPhe		
Db	2229	GATCGCTCGTTTTAAACCGCAACGGGGCTATATCTTTAGGTATACACCGGATGATTACT	2288							
Qy	805	lProTyrArgLeuArgPheAlaGlyAspGlnSerIleArgGlyTyrAlaHisAs	825							

polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 2342 BP; 453 A; 738 C; 716 G; 435 T; 0 other;

Alignment Scores:

Pred. NO.:	3.78e-05	Length:	2342
Score:	182.00	Matches:	139
Percent Similarity:	34.85%	Conservative:	91
Best Local Similarity:	21.06%	Mismatches:	243
Query Match:	3.85%	Indels:	188
DB:	23	Gaps:	25

US-09-914-168-2 (1-919) x AAS89745 (1-2342)

Qy	304	GlyAspValPheHisHis	-----GlyLeuTyrGluThrLysLysAsnLeu	318
Db	361	GGCGAGTTCTCCATACCTACAGACCCACGAGGAGAAAAACCAATGAATCAACCGTTA		320
Qy	319	IleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAspArgSerVal		338
Db	421	CTTTCCGTCGAATACCTGACCCACCTTTACGCCGCGGGCAAA		462
Qy	339	AspValIleLeuProAspAsnThrAlaAspValSer-LeuIleTyrAspThrGlyThrGlu		358
Db	463	-----GGCTTTACGATGTCTCTTTGATTTATGCGCGGGGG		499
Qy	358	nTyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrTh		378
Db	500	-----AAGTGTGGGCATTTGTCGGG		519
Qy	378	rAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeu		395
Db	520	GAATCCG-----GCTCCGGG--AAGACCAGCTGCTGGAAGTCGATCTCCGCGCGCTGCAC		572
Qy	395	rValAsnMetGlyGluAla---TyrAsnLeuGlnAlaValArgAlaLeuSer		411
Db	573	GCCACGACGCGGGAAATTCACACGAGAACCGTTCGCTGATGCAATCAGCAGCGGCCCA		632
Qy	412	---AsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhePr		430
Db	633	CCGCGCTCGCGCTGCTGCGTACCGAATGG---GGCGTGGTGCATCAGCATCCACCG		689
Qy	430	oGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerSerArgTh		450
Db	690	CTTCGCGCGCCAGGTCTCGCAGCGCGGCAATATCGCGAGCGCGCTGATGGCGACCGGGCG		749
Qy	450	rGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValIleLeuTh		470
Db	750	ACGTCAATTACGGCGATATTCGTGCCACCGCGCAGAGATGCTGGAAGAGGTGCAGATCC		809
Qy	470	rAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAs		490
Db	810	CGCCAACCGGATCGACGACCTG---CCGACACACCTTTTCCGCGGCT-----ATCGACA		860

Db 1749 CGGTGATTCGACAAAGTACCGCGGATCTGCGTTCTTCGCGGGGGGACCGCAGTAT 1808
Qy 819 eArgGlyTyrAlaHisAsp-SerLeuSerProIleSerAspLysGly-TyrLeuThrGly 838
Db 1809 TCGTGGCTACCAATACCAAAATCTATCGTCCGGAATATACCAACGCTGACACTGAAAGG 1868
Qy 839 GlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArg 858
Db 1869 GCCTCGAAGTTGATTAACCGGATCGTGGATACCACTACCAACGTACCGGAAATGTACC 1928
Qy 859 LeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLys 878
Db 1929 AAGTGGCAGTTCGGGAA----- 1946
Qy 879 IleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArgVal 896
Db 1947 -----TGGCTGGCGCCC---GGCAGGTGGCGCTG 1973
RESULT 39
AAS93088
ID AAS93088 standard; cDNA; 1916 BP.
AC AAS93088;
DT 13-FEB-2002 (first entry)
XX DNA encoding novel human diagnostic protein #28892.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI: 2001-639362/73.
XX P-PSDB; ABG28901.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX Claim 1: SEQ ID No 28892; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1916 BP; 349 A; 619 C; 582 G; 366 T; 0 other; 1
Alignment Scores:
Pred. No.: 3 51e-05 Length: 1916
Score: 181.00 Matches: 110
Percent Similarity: 34.00% Conservative: 61
Best Local Similarity: 21.87% Mismatches: 166
Query Match: 3.83% Indels: 166
DB: 23 Gaps: 20
US-09-914-168-2 (1-919) x AAS93088 (1-1916)
Qy 458 SerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIle 477
Db 349 GCCACCGCGCAGAAAGTGGCTGGAAGAGGTGGAGATTCGCCCAACCGGATCGACGACCTG 408
Qy 478 SerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnValAlaAla 497
Db 409 ---CCGACACACCTTTTCCGGCGGT-----ATCGACGACGCTTTCAGATT----- 450
Qy 498 LysAlaArgHisLeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAsp 517
Db 451 ---GCCCGCAACCTGTGTGACGATCCG---AAGCTGTGTATTATGATGAACCGACCGCG 504
Qy 518 GlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAla 537
Db 505 GGGCTGGATGTCTCGGTGCGAGCGCCCTGCTCGACCTGCTCGCGCGCTG----- 555
Qy 538 IleLeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAla 557
Db 556 -----GTGGTGGAGCTG-----AACCTCGCG 576
Qy 558 AsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheVal 577
Db 577 -----GTGGTGTATTGTA 588
Qy 578 AlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrProGlySerAspThrGlyThr 597
Db 589 ACCTCTTCCCGACAGACGAGAC-----TGGGGCTTTTACCCACGAGCCT 630
Qy 598 ArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAla 617
Db 631 CGCCTCGCGCGCGCTTC-----CCAGATCCACTCACTCGTCC 669
Qy 618 GluLeuArgLeuSerGluAspLysLys----- 626
Db 670 CGGAGCGCGCAACCGCGCGCCCATCCCGCGCGGAGCTGCCACACCCAGC 729
Qy 627 -----GlyValLysLeuTyrAlaThrLysProLeuSer 637
Db 730 GCACCCACAGTGGCGCGCGGAGATGCGTCAAGCTC---GTCCCGCGCGCGCGCGC 786
Qy 638 HisProLeuAsnAspGlnLeu-----ArgAlaThrLeuGlyTyrGlnGlnGlu--- 653
Db 787 GTGACCCGAATGAGCGCGCGACCTTCGACTCAGTAACTTTGGGGGTTCACACGAGCCCC 846
Qy 653 ----- 653
Db 847 GGGGCGTGTATTGGGACCCCAATTCGACTTTTCCGACCTTCGACTTTCAGCTATATAAT 906
Qy 654 -----ValPheGlyHisSerThrAsnGlyPhe--- 662
Db 907 GCGGTGCTGAAGAATCCACTGGAAACAATATTATTGGT-----GCAGCGCGTTTAAAG 960
Qy 663 -----AspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIle 677
Db 961 CGCACTGACCTGAACGATACCGAATCTGACTCCACTACGCTGGTGGCTTCTCGCTACTGG 1020

Qy	639	ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer	658
Db	1856	TACGATAACTCTTAAAGTGATACACTCTCTAACTATAAGCGTACGACTTATGGAAGTAAT	1915
Qy	659	--ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis-----	671
Db	1916	GTTACTTTAGGTTTCCCTGTAANTGAAATAACTCCTATTATGTAGATTAGGCCATACC	1975
Qy	672	-----GluIleSerArgSerIleIleGlnAsnGlyGlyTyrTrpAsnArgThr-----	686
Db	1976	TATAATAAAATTAGTAACCTTTGCTCTAGAA-----TATAACCGTAAATTATATATT	2026
Qy	687	TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGln	706
Db	2027	CANTCAATGAAATTTAAGGTAAATGCGATTAAACAAAT-----	2065
Qy	707	AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGlyVal	726
Db	2066	-----GACTTTGATTTCTTTTGGTGG-----	2089
Qy	727	AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg	746
Db	2090	-----AACTATACAGCTTAAATAGAGCGTATTTCCCAACT	2125
Qy	747	TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet	762
Db	2126	MAAGGGTTAAAGCAAGCTTGTGGCAGCTTACAATTCACAGTTCTGATACACAATAC	2185
Qy	763	AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer	782
Db	2186	TACAAACTAAGTGCAGATGTACAGGGTTTCTACCCATTA-----	2224
Qy	783	AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp	800
Db	2225	GACAGAGTACACTCTGGGTTGTACTGCAAAAGCATCTGCAGGATATGCCAATGGTTTT	2284
Qy	801	AsnPheAsnHisValProTyrArgLeuArgPheAlaGlyGlyAspGlnSerIleArg	820
Db	2285	GGAAACAAGCGTTTACCGTTCTATCAAACTTATACAGCGGTGGCATTTGTTTCATTACGC	2344
Qy	821	GlyTyrAlaHisAspSerLeuSerPro-----	829
Db	2345	GGTTTGTCTTGTGTAGCATTTGGGCTTACGCATTTATCAGGTCAAAATAATAAATTT	2404
Qy	830	-----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAla	847
Db	2405	AATAAGATAAGTTCTGAT-----GTGATTGGTGGTAATGCAATCGCTACAGTAGCGCA	2458
Qy	848	Glu-----TyrAsnTyrGluPheMetLysAsp-----LeuArgLeuAla	860
Db	2459	GAGTTAATGTGCCAACTCCATTTTGTGAGTGATAGAGTCAAAATACAGTCCGAACTCC	2518
Qy	861	ValPheGlyAspIleGlyAsnAlaTyr-----	869
Db	2519	CTATTGTGTAGCGCGCAAGTGTTCGATACTAAATGGAATCAGATAAAATGGATTA	2578
Qy	870	-----AspLysGlyPheThrAsnAspThrLysIleGly	880
Db	2579	GAGAGCAATGCTTGAAGACCTTACCCGATTTATGGCAAAATCAAGCCGTACTCGCGCTCT	2638
Qy	881	AlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThr	900
Db	2639	ACAGGTGCGGATTCCAATGGCAATCTCTAGTGGACCACTGATTTCTTATGCTAAA	2698
Qy	901	GlyValLys-----GluGluGlyAsnProIleLysLeuHisPheIleGlyThrPro	918
Db	2699	CCAAITAAAAAATATGAAATCATGATGTCGAACAGTTCCAATTTAGTAGTTGGGGTTCT	2758
Qy	919	phe 919	
Db	2759	ttc 2761	

Tue May 6 09:37:17 2003

us-09-914-168-2.rng

Page 64

Search completed: May 6, 2003, 01:20:46
Job time : 949 secs

1

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 5, 2003, 23:56:55 ; Search time 60 Seconds
(without alignments)
4697.266 Million cell updates/sec

Title: US-09-914-168-2

Perfect score: 4727

Sequence: 1 MSPXVLFANRSPFVALAAY.....TGKKEGNPIKLHFFIGTFP 919

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09914168/runat_28042003_151615_6807/app_query.fasta_1.1095
-DB-Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosu62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09914168 @CGN_1.1.32 @runat_28042003_151615_6807 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	181	3.8	2989	3	US-08-433-522A-9
2	181	3.8	2989	3	US-09-135-166-9
3	181	3.8	2989	4	US-08-942-046-9
4	178.5	3.8	2974	3	US-08-433-522A-7
5	178.5	3.8	2974	3	US-09-135-166-7
6	178.5	3.8	2974	4	US-08-942-046-7
7	172	3.6	2949	3	US-08-433-522A-1
8	172	3.6	2949	3	US-09-135-166-1
9	172	3.6	2949	4	US-08-942-046-1
10	172	3.6	2950	3	US-08-433-522A-5
11	172	3.6	2950	3	US-09-135-166-5
12	172	3.6	2950	4	US-08-942-046-5

13	172	3.6	2984	3	US-08-433-522A-3	Sequence 3, Appli
14	172	3.6	2984	3	US-09-135-166-3	Sequence 3, Appli
15	172	3.6	2984	4	US-08-942-046-3	Sequence 3, Appli
16	172	3.6	2987	3	US-08-433-522A-55	Sequence 55, Appl
17	172	3.6	2987	3	US-09-135-166-55	Sequence 55, Appl
18	172	3.6	2987	4	US-08-942-046-55	Sequence 55, Appl
19	159.5	3.4	3957	1	US-07-689-008-5	Sequence 5, Appli
20	159.5	3.4	9540	1	US-07-689-008-1	Sequence 1, Appli
21	137.5	2.9	13121	4	US-08-961-527-126	Sequence 126, App
22	136.5	2.9	38584	4	US-09-453-702B-50	Sequence 50, Appl
23	136	2.9	5976	4	US-08-621-944A-2	Sequence 2, Appli
24	136	2.9	5976	4	US-08-945-567D-2	Sequence 2, Appli
25	136	2.9	6973	1	US-08-478-370-1	Sequence 1, Appli
26	136	2.9	6973	4	US-08-483-855C-1	Sequence 1, Appli
27	136	2.9	6973	4	US-08-621-944A-1	Sequence 1, Appli
28	136	2.9	6973	4	US-08-945-567D-1	Sequence 1, Appli
29	136	2.9	6973	4	US-08-431-718C-1	Sequence 1, Appli
c 30	136	2.9	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c 31	135	2.9	5629	4	US-09-453-702B-243	Sequence 243, App
c 32	135	2.9	30549	4	US-09-134-001C-322	Sequence 322, App
c 33	134.5	2.8	9377	4	US-09-221-017B-1002	Sequence 1002, Ap
34	134	2.8	3121	3	US-08-961-083-117	Sequence 117, App
35	132.5	2.8	32207	2	US-08-770-379-20	Sequence 20, Appl
36	132.5	2.8	32207	4	US-08-757-669A-20	Sequence 20, Appl
37	132.5	2.8	32207	4	US-09-230-371A-20	Sequence 20, Appl
38	131.5	2.8	4411529	4	US-09-103-840A-1	Sequence 1, Appli
39	131	2.8	11091	4	US-09-134-001C-2243	Sequence 2243, Ap
40	130.5	2.8	16836	4	US-09-147-236-1	Sequence 1, Appli
41	130.5	2.8	16836	4	US-09-147-236-10	Sequence 10, Appl
c 42	130	2.8	4411529	4	US-09-103-840A-1	Sequence 1, Appli
43	129	2.7	3758	3	US-08-323-477-1	Sequence 1, Appli
44	129	2.7	14736	4	US-08-961-527-171	Sequence 171, App
45	128.5	2.7	2028	4	US-09-134-001C-1710	Sequence 1710, Ap

ALIGNMENTS

RESULT 1

US-08-433-522A-9
; Sequence 9, Application US/08433522A
; Patent No. 6013514
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: SIA, Dwo Yuan Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6TH Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/433-522A
; FILING DATE: 12-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-434 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 9:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2989 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 390..2768
; US-08-433-522A-9
```

Alignment Scores:

```
Pred. No.: 2,11e-09 Length: 2989
Score: 181.00 Matches: 172
Percent Similarity: 35.57% Conservative: 141
Best Local Similarity: 19.55% Mismatches: 345
Query Match: 3.83% Indels: 222
DB: 3 Gaps: 44
```

US-09-914-168-2 (1-919) x US-08-433-522A-9 (1-2989)

```
Qy 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 378 AGGATCAATCGATGAAACAACTCTAATCGCAAGTTTATTATTCGGTGGCAGCAACGACT 437
Qy 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAlaAsnLysAlaAlaLeuGluAsp 218
Db 438 GTGTTTCGCGCA-----CCCTTTGTGCCAAAGATATTTCGTGTGGATGGT 482
Qy 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 483 GTTCAAGGTGACTTTAGAACACAAATCCGAGCAAGTTTACCT-----524
Qy 239 LeuValAlaAlaArgAlaValGlyTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 525 GTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 581
Qy 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 582 TTCGTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 641
Qy 276 AspTyrArgAla-----ValGluValArgGlyGlyGluGlyAla---Asp 288
Db 642 AGCTTTGTGGCTAAATCGATCTTTCAGATGTTTAAATCAAAGGTAACTCTGTTATTCC 701
Qy 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 702 ACTAAGCAGCTTAACAAACAACTAGATGCTAACCGGTTTAAAGTTGGCGATGTTTA--- 758
Qy 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 759 -----ATTCCAGAAAAATTAAATGAATTTGCCAAAAGTGTAAAAGAGCAGCACTATG 809
Qy 329 PheAspGlyArgTyrProLeuAspArgSerValAspValIle-----LeuProAspAsn 345
Db 810 AGTGTAGTCTGCTAT---AAGCGAACCGTGTGAACCTATTGTCAATACGCTGCCAAATAAT 866
Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 867 CGTCTGNAATTTTAATTCAAATCAATGAAGATGATTAAGCAAAATTTGGCATCATTA 926
Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
Db 927 TTC-----AAGGGGAACGAA---TCTGTGTAGTACAGTACATTTACAAGAA 968
Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr-----402
Db 969 CAAATGGAATATTACAACCTGATTCTTGTGGAAATTTATGGGGAATAAATTTGAAGGTGCG 1028
Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 1029 CAATTCCGAGAAGATCTGCAGCAATTCGT-----GATTAT 1064
```

```
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 1065 TATTTAAATAATGCTATGCCAAAGCACAATACTAAACCGGATGTTTACGTAATATCAT 1124
Qy 439 GlnValSerPheGluGlnSerSerSerSerArgThrGluProAlaGlnValAsp----- 456
Db 1125 GAAAAACAAAAGTTAATTAACCATTTGATGTAATGAAGGTTTACAGTATGACCTTCGT 1184
Qy 457 -----GluSerThrLeuGluProValIleGlu 465
Db 1185 AGTGCACGCATTTAGTAAATCTGGGAGGTATGTCGCCGAGCTTGAACCTTTACTTTCA 1244
Qy 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484
Db 1245 GCATTACATTTAATATGATACTTCCCGCGTAGTGATATTCAGATGTAGAA----- 1295
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaAaGHisLeuTyrAsp 504
Db 1296 ---AATGCAATTAAAGCAAACTTTGGGGAACGAGGTTCACGTAACACACAGTAAATCT 1352
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
Db 1353 GTACCTGATTTTGACGATGCAAAATAAACATTAACGATAACCTTTGTTGTTGATGCTGGA 1412
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
Db 1413 CGAGCTTTAACTGTTCGCCAACTTTCGCTTTGAAGGAATAATCTCTGCTGATGACTACT 1472
Qy 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
Db 1473 TTACGTCAGGAAATGCCAACACAGGAAGAACTTGGTATATTAATTCACAAATTTAGTTGAGTTA 1532
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
Db 1533 GGAATAATTCGCTTAGATCGTACAGGTTTCTCGAAACAGTTTGAACCAAGTAAATGATCCT 1592
Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
Db 1593 ATCAATGGTAGCAATGATGAAGTGGATGCTCGTATATATAAAGTCAAA----- 1637
Qy 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592
Db 1638 -----GAAGCTAACACGGGTAGTATCAACTTTTGGTATTTGGTTACGGT 1679
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608
Db 1680 ACACAGAGTGGTATCAGTTATCAACCAAGTATTAACCAAGATAATTTCTTGGGAACAGGG 1739
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 1740 GCGCAGCTAAGTATAGCTGGTACGAAAAAATGATTATGTCACGAGTGTCAATTTGGGTAT 1799
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
Db 1800 ACCGAACCTATTTTACTAAAGATGGGTAACTCTTGGTGGAAATATTTCTTTGAAANAC 1859
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658
Db 1860 TACCATAACTCTAAAGTGATACATCCTCTAACTATAAGCGTACGACTTATGGAAGTAAT 1919
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671
Db 1920 GTTACTTGTAGTTTCCCTGTAAATGAAATAAATGAAATGAAATGAAATGAAATGAAATGAA 1979
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThr----- 686
Db 1980 TATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2030
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGln 706
Db 2031 CAATCAATGAATTTAAGGTAATGGCATTAAACAAAT-----2069
```

```
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
Db 2070 -----GACTTTGATTTTCTTTGGTTGG----- 2093
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnPrometArgGlyTyrArgGlnArg 746
Db 2094 -----AATATAACAGCCTTAATAGAGCGCTATTTCCCAACT 2129
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 2130 AAGGGGTTAAAGCAAGTCTTGGTGGAGGAGTACTATTCAGGTTCTGATAACAATAC 2189
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 2190 TACAACTAAGTCAGATGATACAGGGTTCTACCACTA----- 2228
Qy 783 AsnArgAlaHisGln-----MetThrGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 2229 GACAGAGATCACCGCTGGTGTGATCTGCAAAAGCATCTCGAGGATGCAAAATGGTTT 2288
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820
Db 2289 GGAACAAGCGTTTACCGTCTTATCAAACTTATACAGCGGTGGCATTTGTCATTAGC 2348
Qy 821 GlyTyrAlaHisAspSerLeuSerPro-----IleSerAspLysGlyTyrLeuThr 837
Db 2349 GGTTCGTTATGTTAGTATTGGGCTTAATGCAATTTATGCGCAACATGGTAATGGTACT 2408
Qy 838 -----GlyGlyGlnValLeuAlaValGlyThrAlaGlu 848
Db 2409 TTTAATAGATAAGTCTCGATGTGTTGATGTAATGCAATCACAACGCGAGTGCAGAA 2468
Qy 849 -----TyrAsnTyrGluPheMetLysAsp-----LeuArgLeuAlaVal 861
Db 2469 CTTATTGTACCAACTCCATTGTTGAGTGATAAAAGCAAAATACAGTCCGAACCTCCCTA 2528
Qy 862 PheGlyAspIleGlyAsnAlaTyr----- 869
Db 2529 TTTTGTGATGCGCAAGTGTGTTGTAATACTAAATGGAATCAGATAAAATGGATTAGAG 2588
Qy 870 -----AspLysGlyPheThrAsnAspThrLysIleGlyAla 881
Db 2589 AGCAAGGTCTTGAAGACTTACCTGATTATGCGCAATCAACCGGATTCGGCGCTCTACA 2648
Qy 882 GlyValGlyValArgTyrPalaSerProValGlyGlnValArgValAspValAlaThrGly 901
Db 2649 GGTCTCGATTCCAATGCAATCTCCTATTGGACCATTTGTTATTTCTTATGCTAAACCA 2708
Qy 902 ValLys-----GluGluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
Db 2709 ATTAATAAATATGAAATGATGATGTCGAACAGTTCCCAATTTAGTATTGGGGGCTCTTTC 2768

RESULT 2
US-09-135-166-9
; Sequence 9, Application US/09135166
; Patent No. 6083743
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: SIA, Dwo Yuan Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6TH Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/135,166
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/433,522
; FILING DATE: 12-SEP-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-829 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2989 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 390..2768
; US-09-135-166-9

Alignment Scores:
Pred. No.: 2,11e-09 Length: 2989
Score: 181.00 Matches: 172
Percent Similarity: 35.57% Conservative: 141
Best Local Similarity: 19.55% Mismatches: 345
Query Match: 3.83% Indels: 222
DB: 3 Gaps: 44

US-09-914-168-2 (1-919) x US-09-135-166-9 (1-2989)
Qy 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 378 AGGATACAATCGATGAAACAACTTCTAATCGCAAGTTTATTATTCGGTGGCACAACGACT 437
Qy 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp 218
Db 438 GTGTTGCCGCA-----CCTTTTGTGCCAAAGATATTCTGTGGATGGT 482
Qy 219 IleThrGlnGluSerAlaMetAspLysSerIleProArgLeuArgGlnThrAla 238
Db 483 GTTCAAGGTGACTTAGAACAAACAATCGAGCAAGTTTACCT----- 524
Qy 239 LeuValAlaAlaArgAlaValGlyTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 525 GTTCGTGCTGCTGCGGTGCTGACTGACATGATG---GCTAATATTGTCGCTCTTTA 581
Qy 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 582 TTCGTAAGTGGTCGATTCGATGTGTGAACGCGCATCAAGAGCGATGCTGCTGTTGTT 641
Qy 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
Db 642 AGCGTTGTGGCTAAATCGATCATTTTCAGATGCTTAAATCAAAAGGTAACCTCTGTTATCCC 701
Qy 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 702 ACTGAAGCACTTAAACAAACTTAGATGCTAACGCGTTTAAAGTTGGCGATGTTTAA--- 758
Qy 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 759 -----ATTGAGAAAAAATAATGAATTTGCCAAAAGTGTAAAGAGACACTATGCA 809
```



```

Qy 862 PheGlyAspIleGlyAsnAlaTyr----- 866
    ||| ||| :: ::
Db 2529 TTTCTTGATCGGCAAGTGTTCGGAATACTAAATGCAAAATCAGATAAAATGCATTAGAG 2588
Qy 870 -----AspLysGlyPheThrAsnAspThrLysIleGlyAla 881
    ||| ||| ::::: ::
Db 2589 AGCAAGTCTTGAAAGACTTTACTGTATGTCGCAAAATCAAGCGGATTCGGCGCTCTACA 2648
Qy 882 GlyValGlyValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThrGly 901
    ||||| ||| ::||| ::||| ::|
Db 2649 GGTCTCGGATCCAAATGGCAATCTCTATTGGACCAATGCTATTTCTTATGCTAAACCA 2708
Qy 902 ValLys-----GluGluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
    ::||| ||| :: ::|
Db 2709 ATTAATAAATATGAAATATGATGATGCGAACAGATTCCAATTTAGTATTGGGGCGCTCTTTC 2768

RESULT 4
US-08-433-522A-7
: Sequence 7, Application US/08433522A
: Patent No. 6013514
: GENERAL INFORMATION:
: APPLICANT: CHONG, Pele
: APPLICANT: THOMAS, Wayne
: APPLICANT: YANG, Yan Ping
: APPLICANT: LOOSMORE, Sheena
: APPLICANT: SIA, Dwo Yuan Charles
: APPLICANT: KLEIN, Michel
: TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
: NUMBER OF SEQUENCES: 55
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: 6TH Floor, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/433,522A
: FILING DATE: 12-SEP-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: STEWART, Michael I
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-434 MIS:jb
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2974 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 386..2761
: US-08-433-522A-7

```

Alignment Scores:	3.95e-09	Length:	2974
Pred. No.:	178.50	Matches:	173
Score:	3.30%	Conservative:	138
Percent Similarity:	19.64%	Mismatches:	345
Best Local Similarity:	3.78%	Indels:	225
Query Match:	3	Gaps:	44
DB:			

US-09-914-168-2 (1-919) x US-08-433-522A-7 (1-2974)

```
Qy 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 374 AGGATACAATCGCAAAACCTTCTAATCGCAAGTTTATTATTTCGGTACGACACGACT 433
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLeuGluAsp 218
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 434 GTGTTTCCGCA-----CCTTTGTGTGGCAAAAGATATTTCGTGTGATGGT 478
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 479 GTTCAAGTGACTTAGAACACAATCCGACAGATTACCT----- 520
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 521 GTTCGTGCGGTGACGGTGTGACTGACAAATGATGFG---GCTAATATTGTCGGTCCTTTA 577
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 578 TTCGTAAGTGTGCGATTGCGATGTTGAAAGCGCATCAAGAGCGGATGCTTGTGTT 637
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 276 AspTyrArgAla-----ValGluValAlaArgGlyGluGly---AlaAsp 288
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 638 AGCGTTGTGGCTAATTCGATCATTTACATGTTAAATCAAGGTAACCTCTATTATTCGA 697
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 698 CCTGAAGCACTAAAACAAACTTAGATGCTAAAGGGTTAAAGTTGGCGATATTTA--- 754
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 755 -----ATTCGAGAAAAATTAATGAATTTGCCAAAGTGTAAAAGACACATATCCA 805
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 329 PheAspArgTyrLeuAspArgSerValAspValIle-----LeuProAspAsn 345
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 806 AGTGTAGTCCGTAT---AACGCAACCGTTGAACCTATTGTCAATACCGTACCAATAAT 862
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 863 CGCGCTGAATTTTAATTCAAATCAATCAAGATGATAAGCAAAATGGCATCATTAAT 922
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 923 TTC-----AAGGGAACGAA---TCTGTTAGTACGAGTACATTACAAGAA 964
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 965 CAATGGGAATTACAACCTGATCTCTGGTGGAAATTTATCGGGAATAAATTTGAAGTCCG 1024
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1025 CAATTGCAAGAAAGATTTGCGAGCAATTCGT-----CATTAT 1060
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1061 TATTTAATAATGCGTATGCCAAAGCACAAATCACTAAAGCGGATGTTTCAGCTAAATGAT 1120
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1121 GAAAAACAAGTAATGTAAACCATTTGATGTAATGAAGGTTTACAGTATGACCTTCGT 1180
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 457 -----GluSerThrLeuGluProValIleGlu 465
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1181 AGTGCACCATATAGGTAATCTGGGAGGTATGTCGCGGAGCTTGAACCTTACTTCA 1240
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1241 GCATTACATTTAAATGATACCTTTCCGCGGTAGTGATATTGCAGATGTAGAA----- 1291
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1292 ---AATGCAATTAAGCAAAACTTGGGGAACGAGGTTACGGTAAACACACAGTAAATCT 1348
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
```

```
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspGlyValAsn 520
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1349 GTACCTGATTTTGACGATCAATAAACATTTAGCGATAACCTTTCTTGTGTATGCTGGA 1408
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1409 CGACGTTTAACTGTTTCACCAACTTCGCTTTGAAGGAATAACGCTTCTCTGTATGACT 1468
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1469 TTACGTCAGAAATCGCCACACAAGAAGAACTTGTGTATAATTCACAATTAGTTAGTTA 1528
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1529 GGAATAATTCGCTTAGATCGTACAGGTTTCTCGAAACAGATTGAAACCGAATTTGATCCT 1588
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1589 ATCAATGCTAGCAATGATGAAGTGGATGCTGTATATAAAGTCAAA----- 1633
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTyrGly 592
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1634 -----GAACGTAACACGGGTAGTATCACTTTGGTATTGGTTACGGT 1675
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1676 ACAGAGAGTGGTATTAGTTATCAAGCAAGTGTCAACAAGATAAATTTCTTGGCAACAGGG 1735
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1736 GCGGCAGTAAGTATAGCTGTACGAAATAATGATTATGGTACGAGTGTCAATTTGGGTTAT 1795
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1796 ACCGAGCCCTATTTTACTAAAGATGGTGAAGTCTTGGTGGAAATCTTTCTTTGAAAC 1855
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1856 TACGATAACTTAAAGTGATACATCTCTAATACTATTAAGCGTACGACTTATGGAAGTAAT 1915
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1916 GTTACTTTAGGTTTCCCTGTAATGAAATAACTCTCTATATGTAGGATGAGCCCATACC 1975
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThr----- 686
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 1976 TATAATAAATTAGTAACCTTTCCTCTAGAA-----TATAACCGCTAATTTATATATT 2026
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGln 706
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 2027 CAATCAATGAAATTTAAAGGTAATGGCATTAACAACAAAT----- 2065
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGlyVal 726
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 2066 -----GACTTTGATTTTCTTTTGGTGG----- 2089
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 2090 -----AACTATAACAGCCTTAATAGAGCGCTATTTCCCACT 2125
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 2126 AAAGGGTTAAAGCAAGTCTTGTGGACGAGTTTACAATTCAGAGTCTGTATAACAATAC 2185
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 2186 TACAACAACTAAGTCAGATGTACAGGGTCTCTACCCATTA----- 2224
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 2225 GACAGAGATCACCTCTGGGTTGTATCTGCAAAAGCATCTGCAGGATATGCAAAATGTTT 2284
    ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
```

```
QY 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820
Db 2285 GGAACAAGCGTTTCCTTATCAAACTTATACACGCGGTGGCATTTGGTTCAATACGC 2344
QY 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2345 GGTTTTCCTTATCGTACGATGGCGCTAACGCAATTTATCAAGGTCAAAATAATAAATTT 2404
QY 830 -----IleSerAspLysGlyTyrLeuThrGlyGlnValLeuAlaValGlyThrAla 847
Db 2405 AATAAGATAAGTCTGAT-----GTGATTGGTGGTAATCAATCGCTACAGCTAGCGCA 2458
QY 848 Glu-----TyrAsnTyrGluPheMetLysAsp-----LeuArgLeuAla 860
Db 2459 GAGTTAATTGCGCAACTCCATTTGTGAGTGATAAGAGTCAAAATACAGTCCGGAACCTCC 2518
QY 861 ValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2519 CTATTTGTTGATCGGCAAGTGTGGAATCTAAATGGAATCAGATAAAATGGAATTA 2578
QY 870 -----AspLysGlyPheThrAsnAspThrLysIleGly 880
Db 2579 GAGAGCAATGCTTTGAAGACTTACCGGATTTATGGCAAAATCAAGCGGTACTCGCGCTCT 2638
QY 881 AlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThr 900
Db 2639 ACAGTGTGGATTCCCAATGGCAATCTCCTAGTGGACCACTGCTATTTCTTATGCTAAA 2698
QY 901 GlyValLys-----GluGluGlyAsnProIleLysLeuHisPheIleGlyThrPro 918
Db 2699 CCAATTAAAAAATATGAAATGATGATGTCGAACAGTTCCAATTTAGTATTGGGGTTCT 2758
QY 919 Phe 919
Db 2759 TTC 2761

RESULT 5
US-09-135-166-7
; Sequence 7, Application US/09135166
; Patent No. 6083743
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: SIA, Dwo Yuan Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6TH Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/135,166
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/433,522
; FILING DATE: 12-SEP-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-829 MIS:jb
```

```
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2974 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 386..2761
; US-09-135-166-7

Alignment Scores:
Pred. No.: 3,95e-09 Length: 2974
Score: 178.50 Matches: 173
Percent Similarity: 35.30% Conservative: 138
Best Local Similarity: 19.64% Mismatches: 345
Query Match: 3.78% Indels: 225
DB: Gaps: 44

US-09-914-168-2 (1-919) x US-09-135-166-7 (1-2974)
QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 374 AGGATACAACTCGATGATAAAAAAAGCTTCTAATCGCAAGTTTATTTCGGTACGACACGACT 433
QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLeuGluAsp 218
Db 434 GTGTTTGGCGCA-----CCTTTGTGGCAAAAGATATTTCGTGTGGATGT 478
QY 219 IleThrGlnGlnSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 479 GTTCAAGGTGACTTAGAACACAAATCCGAGCAAGTTTACCT----- 520
QY 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 521 GTTCGTGGCGGTGAGCGTGTGACTGACATGATGTG---GCTAATAATTGCGCTCTTTA 577
QY 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 578 TTCGTAAGTGGTGGATTCGATGATGGAAGCGCATCAAGAGCGCATGTCTTGTGT 637
QY 276 AspTyrArgAla-----ValGluValArgGlyGluGly---AlaAsp 288
Db 638 AGCGTTGTGCTAAATCGATCATTTTCAGATGTTAAATCAAAAGTAACCTATTATCCA 697
QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 698 CCGTAAGCACTAAACAAACTTAGATGCTAACGGGTTTAAAGTTGGCGATATTTTA--- 754
QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 755 -----ATTCGACAAAAATTAATGAATTTGCCCAAGTCTAAAGAGCACTATGCA 805
QY 329 PheAspGlyArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsn 345
Db 806 AGTGTAGGTGCGTAT---AACGCAACCGTTGAACCTATTGTCAATACGCTACCAATAAT 862
QY 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 863 CGCGCTGAAATTTTAAATCAATGAAGATGATAAGCCCAAAATTTGCATCATTAAC 922
QY 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
Db 923 TTC-----AAGGGGAACGAA---TCTGTAGTAGCAGTACATTACAAGAA 964
QY 386 LysArgGluLeuGluGlnLeuThrValAsnMetGlyGluAlaTyr----- 402
Db 965 CAAATGGAATTACAACTCATTTCTTGGTGGAAATTTATGGGAAATAAATTTGAAGGTGCG 1024
```


Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 1025 CAATTCGAGAAAGATTTTCAGGCAATTCG-----GATTAT 1060
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 1061 TATTTAAATAATGCGTATGCGCAAGCACAATACTAAGCGGATGTTTCAGCTAAATGAT 1120
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456
Db 1121 GAAAAACAAAAGTTAATGTAACCATTCATGTAATGAAGGTTTACAGTACGCTTCGT 1180
Qy 457 -----GluSerThrLeuGluProValIleGlu 465
Db 1181 AGTCACGCATATAGGTAATCTGGAGGTATGTCGCGAGCTTGAACCTTTACTTCA 1240
Qy 466 ThrValGluLeuThrAspGlyLeu---LeuMetAspIleSerProIleGluPheSerAla 484
Db 1241 GCATTACATTTAATGATACCTTCGCGCGTAGTGATATTCAGATGTAGAA----- 1291
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
Db 1292 --AATGCAATTAAGCAAACTTGGGCAACGAGGTTACGGTTAACACACAGTAATCT 1348
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspGlyValAsn 520
Db 1349 GTACCTGATTTGACGATGCAAAATAAAACATTTAGCGATAACCTTTGTTGATCGCTGA 1408
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
Db 1409 CGACGTTTAACCTGTTCCACCACTTCGCTTGAAGGAATAACGGTTTCGCTGATAGTACT 1468
Qy 534 ValAlaArgAlaIleLeuProAspIleSer-----GluAsnGluValIleAspLeu 550
Db 1469 TTACGTCAGGAATCGCGCAACAGAGAACTTGGTATAATTCACAATTAGTTAGTTA 1528
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
Db 1529 GGAATAATTCGCTTAGATCGTACAGGTTTCTTCAAGACAGTTGAAACCGAATTCATCCT 1588
Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
Db 1589 ATCAATGCTAGCAATGATGAAGTGGATGCTGTATATAAGTCAAA----- 1633
Qy 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592
Db 1634 -----GAACGTAACACACGCGTAGTATCACTTTGGTATTGGTTACGGT 1675
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608
Db 1676 ACAGAGAGTGTATTAGTTATCAAGCAAGTGTCAACCAAGATATTTCTTGGGAACAGGG 1735
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 1736 CCGCAGTAAAGTATAGCTGCTACGNAATAATGATATGTACGAGTCTCAATTTGGGTTAT 1795
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
Db 1796 ACCGAGCCCTATTTTACTAAAGTGGTAAAGTCTTGGTGGAAATGTTTCTTTTGAAC 1855
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658
Db 1856 TAGGATAACTCTAAAGGTAGTACATCTCTTAACATAAGCGTACGACTATGGAAGTAA 1915
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671
Db 1916 GTTACTTTAGGTTTCCCTGTAATAAGCAATACTCCTATTATGTAGGATTAGGCCATACC 1975
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThr----- 686
Db 1976 TATAATAAAATAGTAACCTTGCTCTAGAA-----TATAACCGTAATTTATATATT 2026
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGln 706

Db 2027 CAATCAATGAAATTTAAAGGTAATGCGATTAACAAAT----- 2065
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGlyVal 726
Db 2066 -----GACTTTGATTTTCTTTTGTGG----- 2089
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
Db 2090 -----AACTATAACAGCCTTAATAGAGGCTATTTCCTCAACT 2125
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 2126 AAAGGGGTTAAAGCAAGTCTTGTGTGACAGGTACAAATTCAGGTTCTGTATACAAATAC 2185
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 2186 TACAACAACTAAGTGCAGATGTACAGGTTTCTACCCATTA----- 2224
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 2225 GACAGAGATCACCCTCTGGGTTGTATCTGCAAAAGCATCTGCAGGATATGCAAAATGGTTT 2284
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820
Db 2285 GGAACAACAGCGTTTACCCTTCTATCAAACTTATACAGCGGTTGGTTCATTACGC 2344
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2345 GGTTCGTTTATGTTAGTACATTTGGCCTACGCATTTATCAAGGTCAAAATAATAATTT 2404
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAla 847
Db 2405 AATAAGATAAGTTCTCAT-----GTGATTGGTGAATGCAATCGCTACAGCTAGCGCA 2458
Qy 848 Glu-----TyrAsnTyrGluPheMetLysAsp-----LeuArgLeuAla 860
Db 2459 GAGTTAATGTGCCAACTCCATTTGTGAGTGATAGAGTCAAAATACAGTCCGAACCTCC 2518
Qy 861 ValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2519 CTATTTGTTGATGCGGCAAGTCTTGGATACTAATAATGGAATCAGATAAAATGGATTA 2578
Qy 870 -----AspLysGlyPheThrAsnAspThrLysIleGly 880
Db 2579 GAGAGCAATGTCTTGAAGACTTACCGGATTTATGGCAAAATCAAGCGTACTCGCGCTCT 2638
Qy 881 AlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThr 900
Db 2639 ACAGGTGTCGGAATCCAATGGCAATCTCTAGTGACCACTCTTCTTTATGTCTAA 2698
Qy 901 GlyValLys-----GluGluGlyAsnProIleLysLeuHisPheIleGlyThrPro 918
Db 2699 CCAATTAATAAATATCAAAATGATGATGTCGAACAGTTTCAATTTAGTATTGGGGTTCT 2758
Qy 919 Phe 919
Db 2759 TTC 2761

RESULT 6

US-08-942-046-7
; Sequence 7, Application US/08942046
; Patent No. 6264954
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: SIA, Dwo Yuan Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney
 STREET: 6TH Floor, 330 University Avenue
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5G 1R7

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/942,046
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/433,522
 FILING DATE: 12-SEP-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: STEWART, Michael I
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1038-732 MIS:jb
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2974 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 386..2761
 US-08-942-046-7

Alignment Scores:
 Pred. No.: 3.95e-09 Length: 2974
 Score: 178.50 Matches: 173
 Percent Similarity: 35.30% Conservative: 138
 Best Local Similarity: 19.64% Mismatches: 345
 Query Match: 3.78% Indels: 225
 DB: 4 Gaps: 44

US-09-914-168-2 (1-919) x US-08-942-046-7 (1-2974)

QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
 Db 374 AGGATCAATCGATGAAACAACTTCTAATCGCAAGTTTATTATTCGGTAGCGCAACGACT 433
 QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLalaLeuGluAsp 218
 Db 434 GTCTTGGCCCA-----CCTTTGTGGCAAAAGATATTCGTGTGGATGGT 478
 QY 219 IleThrGlnGlnSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
 Db 479 GTTCAAGGTGACTTAGAACCAACAAATCCGAGCAAGTTTACCT-----520
 QY 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
 Db 521 GTTCGGTCGGTCAGCGGTGACTGACATGATGTG---GCTAATATTGTCCGCTCTTTA 577
 QY 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
 Db 578 TTCGTAAGTGGTCGATCGATGATGTGAAGCCCATCAAGAGCGCATGTGCTTGTGTT 637
 QY 276 AspTyrArgAla-----ValGluValArgGlyGluGly---AlaAsp 288
 Db 638 AGCGTTGTGCTAATATCGATCATTTTCAGATGTTAAATCAAAGGTAACCTATTATTCCA 697
 QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308

Db 698 CCTGAAGCACATAAAACAACTTAGATGCTAACGGCTTTAAAGCTTGGCGATATTTTA---754
 QY 309 HisGlyLysTyrCluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
 Db 755 -----ATTGAGAGAAAATTAATGAATTTGCCAAAGTGTAAAGAGCATTATGCA 805
 QY 329 PheAspGlyArgTirPleuAspArgSerValAspValIle-----LeuProAspAsn 345
 Db 806 AGTGTAGTTCGCTAT---AACGCCACCGTTGAACCTATTGTCAANTACGCTACCAANAAT 862
 QY 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
 Db 863 CGCGCTGAAATTTTAATTCAAATCAATGAAGATGATAAAGCCCAAAATGGCATCATTAAC 922
 QY 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
 Db 923 TTC-----AAGGGGAACGAA---TCTGTAGTAGCAGTACATTACAAGAA 964
 QY 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr-----402
 Db 965 CAAATGGAATTACAACCTGATTCTTGGTGGAAATTAATGGGAAATAAATTTGAAGGTGCG 1024
 QY 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
 Db 1025 CAATTTCGAGAAAGATTTCAGGCAATTTCGT-----GATTAT 1060
 QY 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
 Db 1061 TATTTAAATAATGGCTATGCCAAGCACAATCACTAAACCGGATGTTTCAGCTAAATGAT 1120
 QY 439 GlnValSerPheGluGlnSerSerSerSerArgThrGluProAlaGlnValAsp-----456
 Db 1121 GAAAAACAAAGTTAATGTACCACTGATGTAATGAAGTTTACAGTATGACCTTCGT 1180
 QY 457 -----GluSerThrLeuCluProValIleGlu 465
 Db 1181 AGTGCACGCATTATAGGTAAATCTGGAGGTATGCTCTGCGAGCTTGAACCTTTACTTTCA 1240
 QY 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484
 Db 1241 GCATTACATTTAAATGATACTTCTCCGCCCTGATGATATTGCACATGCTACAA-----1291
 QY 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
 Db 1292 ---AATGCAATTAAGCAAAACTTGGGAAACGAGGTAGCGTAACACAAACAGTAATTTCT 1348
 QY 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
 Db 1349 GTACCTGATTTTGACGATGCAATAAAACATTAGCCATTAACCTTTGCTTGTGCTGCTGGA 1408
 QY 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
 Db 1409 CGACGTTTAACTGTTCAACCACTTCCTTTGAAGGAATAACCTTTCTGCTGATGACTACT 1468
 QY 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
 Db 1469 TTACGTCAGGAATTCGCCCAACAAAGGAACCTTGTGTATAATTCACAATTAGTTCAGTTA 1528
 QY 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
 Db 1529 GGAAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAACAGTTGAAACCCGAATTCATCCT 1588
 QY 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
 Db 1589 ATCAATGGTAGCAATGATGAAGTGGATGCTGCTATATATAAAGTCAAA-----1633
 QY 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyThrPgly 592
 Db 1634 -----GAACGTAACACGGGTAGTAGTCAACTTTGTTGTTGTTACGCT 1675
 QY 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle-----608

```

Qy  901 GlyValIys-----GluGluGlyAsnProIleLysLeuHisPhePheIleGlyThrPro 918
      ::::: ||| ::::: ::::: ||| :::::
Db  2699 CCAATTAATAATATGAAATGATGATGCGAACACAGTCCCAATTTAGTATTGGGGTTCT 2758

Qy  919 Phe 919
      |||
Db  2759 TTC 2761

RESULT 7
US-08-433-522A-1
; Sequence 1, Application US/08433522A
; Patent No. 6013514
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: SIA, Dwo Yuan Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6TH Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/433.522A
; FILING DATE: 12-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I.
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-434 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2949 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 75..2465
US-08-433-522A-1

Alignment Scores:
Pred. No.: 2.03e-08 Length: 2949
Score: 172.00 Matches: 171
Percent Similarity: 35.21% Conservative: 141
Best Local Similarity: 19.30% Mismatches: 344
Query Match: 3.64% Indels: 230
DB: 3 Gaps: 44

US-09-914-168-2 (1-919) x US-08-433-522A-1 (1-2949)

Qy  179 LysValProArgLeuLysAlaLysPheTyGlnSerGlnSerGlyGlyThrSerAla 198
      ::::: ::::: ||| ::::: ||| :::::
Db  63 AGGTACAAATCGATGAAAGAAACTTCTAATCGCAAGTTTATTATTCGGTACGACACGACT 122

Qy  199 IleGlySerHisGlnLysThrGluProTyGlnAlaAsnIleLysAlaLeuGluAsp 218
      ::::: ::::: ||| ::::: ||| :::::

```

Db 123 GTGTTTCCCGCA-----CCTTTTGGCAAAAGATATTCGTGGTGGT 167
Qy 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 168 GTTCAAGGTGACTTAGAACACAAATCCGAGCAAGTTACT----- 209
Qy 239 LeuValAlaAlaArgAlaValGlyTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 210 GTTCGGTCCGGTCAGCGTGTGACTGACAATGATGTG---GCTAATATGTCCTCTTTA 266
Qy 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 267 TTCGTAAGTGGTCGATTCGATGATGTGAAGCCATCAAGAGCGGATGTGCTTGTGTT 326
Qy 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
Db 327 AGCGTTGTGGCTAAATCGATCAATTCAGATGTTAAATCAAAAGTAACTCTGTATTCC 386
Qy 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 387 ACTGAAGCACTTAACAAACACTTAGTCTAACGGGTTTAAAGTTGGCGATGTTTA--- 443
Qy 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 444 -----ATTCGAGAAAATTAATGAATTTGCCAAAAGTGTAAAGAGCACTATGCA 494
Qy 329 PheAspGlyArgTrpLeuAspArgSerValaspValIle-----LeuProAspAsn 345
Db 495 AGTGAGTGGCGTAT---AACCCACACAGTTGAACCTATTGTCAATACGCTACCAATAAT 551
Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 552 CGCGCTGAAATTTAATCAATCAATGAGATGATAAGCAAAATTTGGCATCATTAAC 611
Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
Db 612 TPTC-----AAGGGGAACGAA---TCTGTTAGTAGCAGTACATTACAGAA 653
Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402
Db 654 CAATGGAAATACACACCTGATCTTGGTGGAAATTTATGGGAAATTAATTTGAAGGTGCG 713
Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 714 CAATTCGAGAAAGATTTCAGTCAATTCGT-----GATTAT 749
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 750 TATTTAAATAATGGCTATGCCAAAGCACAAATTTACTAAACCGGATGTTTCAGCTAAATGAT 809
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456
Db 810 GAAAAACAAAAGTTAATGTACCATTCATGTAATGAAGCTTACAGTATGACCTTCGT 869
Qy 457 -----GluSerThrLeuGluProValIleGlu 465
Db 870 AGTCACGCAATTATAGTAATCTGGAGGTATGCTCGCGAGCTTGAACCTTTACTTTCA 929
Qy 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484
Db 930 GCATTACATTTAATGATCTTCCGCGGTAGTGATATTGCAGATGTAGAA----- 980
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
Db 981 ---AATGCAATTAAGCAAACTTTGGAGACCGCGTGTACGGTAGCCGCAACGGTAATTC 1037
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
Db 1038 GTACCTGATTTTGATGATGCAAAATTAACCATTTAGCGATAACCCCTTGTGTGTGCTGGA 1097
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
Db 1098 CGAGTTTAACGTGTCGCAACATTCGCTTTGAAGGAATAACCGTTCTGCTGATGACCT 1157

Qy 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
Db 1158 TTACGTCAGGAAATGCCCAACAAGAAGAACTTGGTATAATTCACAATTTAGTTGAGTTA 1217
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
Db 1218 GGAAAAATTCGCTTAGCTAGCTACAGGTTCTTCGAACACAGTCCAAACCGAATGATCCT 1277
Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
Db 1278 ATCAATGGTAGTAATGATCAAGTGGATGTCGTATATAAAGTCAAA----- 1322
Qy 575 ValPheValAlaSerAspLysProArgaspGlyGlnIle-----GlyLeuGlyTrpGly 592
Db 1323 -----GAACGTAAACAGCGGTAGTATCAACTTTGGTATTGTTGTTACGCT 1364
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608
Db 1365 ACAGAGAGTGGTATTAGTTATCAAGCAAGTGTAAACAAGATAATTTCTTGGGAACAGGG 1424
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 1425 GCGGCAGTAAGTATAGCTGGTACGAAATAATGATTATGGTACGAGTCTCAATTTGGTTAT 1484
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
Db 1485 ACCGAGCCCTATTTTACTAAAGATCGTGTAGTCTTGGTGGAAATGTTTCTTCTTCAAAAC 1544
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658
Db 1545 TACGATAACTCTAAAGTGATACATCTTAACATAAGCGTACGACTTACGGAAGTAA 1604
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671
Db 1605 GTTACTTTAGGTTTCCCTGTAATGAAATAAATTAACCTTATATGTAGGATTAGTCATACC 1664
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThr----- 686
Db 1665 TATAATAAATTAGTAACCTTTGCTCTAGAA-----TATAACCGTAATTTATATATT 1715
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGln 706
Db 1716 CAATCAATGAAATTTAAAGGTAAATGGCATTAACAAACAAAT----- 1754
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
Db 1755 -----GACTTTGATTTTCTTTTGGTTGG----- 1778
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnPrometArgGlyTyrArgGlnArg 746
Db 1779 -----AACTATAACAGCCTTAATACAGGCTATTTCACCAACT 1814
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 1815 AAAGGGGTTAAAGCAAGCTCTGGTGACGAGTTACTATTCCAGGTTCTGTATAACAATAC 1874
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 1875 TACAAACTAAGTSCAGATGTACAGGTTTCTACCCATTA----- 1913
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 1914 GACAGAGATCACCTCTGGGTTGTATCTGCAAAAGCATCTCGAGGATATGCAATTTGTTT 1973
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyAspGlnSerIleArg 820
Db 1974 GGAACACAGCGTTTACCGTCTCTATCAAACTTATACAGCGCGGTGGCATCGGTTTAC 2033
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2034 GGTTTTGTCTTATGTTAGTATTGGACCAATTTATGCCCAATATATGCCCAATATGTAATG 2093

QY 830 -----IleSerAspLysGlyTyrLeuThrGlyGlnValLeu 842
Db 2094 GGTACTGGTACTTTAAGAAGATAAGTTCTGAT-----GTGATTGGTGGTAATGCAATC 2147
QY 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856
Db 2148 GCTACAGCTAGCGCAGAGTAAATTGTGCCAACTCCCAATTTGTGAGCGATAAGAGCCAAAT 2207
QY 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2208 ACGGTCCGACCTCTCTATTGTGATCGGCAAGTGTGGAACTAAATGGAATCA 2267
QY 870 -----AspLysGlyPheThrAsn 875
Db 2268 GATAAAATGGATTAGAGCGCATGTATTAAGAAAGATTCCCTGATTATGCCAAATCAAGC 2327
QY 876 AspThrLysIleGlyAlaGlyValAlaArgTTPAlaSerProValGlyGlnValAlaArg 895
Db 2328 CQTATTCGGCCCTCTACAGGTCTCGATTCCCAATGGCAATCTCCTATTGGGCCATTGGTA 2387
QY 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
Db 2388 TTCCTCTATGCCCAACCAATTAATAATATGAAATGATGATGTCGAACAGTTCCAATTT 2447
QY 914 PheIleGlyThrProPhe 919
Db 2448 AGTATTGGAGTCTTCTTC 2465

RESULT 8

US-09-135-166-1
; Sequence 1, Application US/09135166
; Patent No. 6083743
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: SIA, Dwo Yuan Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN.
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6TH Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/135,166
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/433,522
; FILING DATE: 12-SEP-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-829 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2949 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 75..2465
US-09-135-166-1
Alignment Scores:
Pred. No.: 2,03e-08 Length: 2949
Score: 172.00 Matches: 171
Percent Similarity: 35.21% Conservative: 141
Best Local Similarity: 19.30% Mismatches: 344
Query Match: 3.64% Indels: 230
DB: Gaps: 44
US-09-914-168-2 (1-919) x US-09-135-166-1 (1-2949)
QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 63 AGGATACAATCGATGATAAAAAAATCTTAATCGCAAGTTTATTATTCGGTAGCACAACGACT 122
QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp 218
Db 123 GTGTTTGGCGCA-----CCTTTGTGCAAAAGATATTCTGTGTGGATGCT 167
QY 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 168 GTTCAAGGTGACTTAGACAACAATAATCCGAGCAAGTTTACT----- 209
QY 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 210 GTTCGTCCGGTCCGCTGCTGACTGACAATGATGTG---GCTAATATTGTCCGCTCTTTA 266
QY 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 267 TTCGTAAGTGGTCGATTCGATGTGAAGCGCATCAAGAGCGGTTAAAGTTGGCGATGTTT 326
QY 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
Db 327 AGCGTTGTGCTAAATCGATCATTTTCAGATGTTTAAAAATCAAGGTAACCTCTGTATTCCC 386
QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 387 ACTGAAGCACTTAAACAAACTTAGTAGCTAACGGGTTTAAAGTTGGCGATGTTTAA--- 443
QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 444 -----ATTCGAGAAAAATTAATGAATTTGCCAAAAGTGTAAAAGACACTATGCA 494
QY 329 PheAspGlyArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsn 345
Db 495 AGTGTAGTGGTGTAT---AACGCAACAGTTGAACCTATTGTCAATAGCTACCAATAAT 551
QY 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 552 CGCGCTGAATTTTAATCAATCAATGAAGATGATGAAGCAAAATGGCATCATTAAC 611
QY 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
Db 612 TTC-----AAGGGGAACGAA---TCCTTAGTAGCAGTACATTACAAGAA 653
QY 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402
Db 654 CAAATGGAATTACAACCTGATTCTTGTGTGAAATTTATGGGAAATAAATTTGAAGTGGC 713
QY 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 714 CAATTCGAGAAAGATTTCGAGTCAATTCGT-----GATTAT 749
QY 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 750 TATTTAAATAATGGCTATGCCAAAGCACAAATTAATAAACGGATGTTTCACTGAATGAT 809

```

QY 439 GlnValSerPheGluGlnSerSerSerSerArgThrGluProAlaGlnValAsp----- 456
Db 810 GAAAAACAAAAGTAATGTAAACCAATTGATGTAATGAAGTTTACAGTATGACCTTCGT 869
QY 457 -----GluSerThrLeuGluProValIleGlu 465
Db 870 AGTCCAGCCATTATAGTAATCTCGGAGGTATGCTCCCGAGCTTGAACCTTTACTTTCA 929
QY 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484
Db 930 GCATTACATTAAATGATGATCTTCGCGCGTAGTGATATGCAGATGTAGAA----- 980
QY 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
Db 981 ---AATGCAATTAAGCAAAATCTTCCGCGTAGTGATATGCAGATGTAGAA----- 980
QY 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
Db 1038 GTACCTGATTTTGATGATGCAATAAACAATAGCGATACCCCTGTTGTTGATCGTGA 1097
QY 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
Db 1098 CGAGCTTTAACTGTTCCGCAACTTCGCTTTGAAGGAATACCGTTTCTGCTGATAGCACT 1157
QY 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
Db 1158 TTACGTGAGGAATGCGCAACACAGNAGGAAGTGTGTAATAATTCACAAATAGTTGAGTTA 1217
QY 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
Db 1218 GGAATAAATTCGTTAGTGTACAGGTCTTCTCGAAACAGTCCGAAACCGAATTCATCCT 1277
QY 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
Db 1278 ATCAATGGTAGTAATGATGAAGTGGATGCTGTATATAAGTCAAA----- 1322
QY 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592
Db 1323 -----CAACGTAAACACGGGTAGTATCAACTTTGGTATGTTGTTACCGT 1364
QY 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608
Db 1365 ACAGAGAGTGGTATTAGTTATCAACCAAGTGTAAACCAAGATAATTTCTTGGGAACAGGG 1424
QY 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 1425 CGCGCAGTAAGTATAGCTGGTACGAAAAATGATTATGTGTACGAGTGTCAATTTGGGTAT 1484
QY 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
Db 1485 ACCGAGCCCTATTACTAAAGATGGTGAAGTCTTGGTGGAAATGTTTCTTTGAAAAAC 1544
QY 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658
Db 1545 TACGATAACTCTAAAGGTATACATCTCTAACTATAGCGTAGCACTTACGGAAGTAAT 1604
QY 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671
Db 1605 GTTACTTTAGGTTTCCCTGTAATAGAAATAAATCTCTATTATGTAGGATTAGGTCAACC 1664
QY 672 -----GluIleSerArgSerIleGlnAsnGlyGlyTyrAsnArgThr----- 686
Db 1665 TATAATAAATAGTAACTTTGCTCTAGAA-----TATAACCGTAATTTATATATT 1715
QY 687 TyrSerLeuArgTyrArgLeuAspLysLysThrGlnAlaProGluThrTrpGln 706
Db 1716 CAATCAATGAAATTAAGGTGAATGGCATTAAACAAAT----- 1754
QY 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGlyVal 726
Db 1755 -----GACTTTGATTTTCTTTGGTTGG----- 1778
QY 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746

```

```

Db 1779 -----AACTATAACAGCCCTTAATAGAGGCTATTTTCCCAACT 1814
QY 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 1815 AAAGGGGTTAAACCAAGTCTTGGTGCACGAGTTACTATTCCAGGTTCTGATAAACAAATAC 1874
QY 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 1875 TACAACAACTAAGTGCAGATGTACAGGGTTTCTACCCATTA----- 1913
QY 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 1914 GACAGATCACCTCTGGGTTGTATCTGCAAAAGCATCTGCAGGATATGCAAAATGGTTT 1973
QY 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820
Db 1974 GCAACAAAGCGTTTACCGTTCTATCAAACTTATACAGCGGGTGGCATCGGTTTCATTACGT 2033
QY 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2034 GGTITGCTTATGGTAGTATTGGACCTTAACGCAATTTATGCCGAATATGGTAATGGTAGT 2093
QY 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842
Db 2094 GGTACTGTACTTTTAAACAAGATAAGTTCTCAT-----GTGATTGGTGGTAATGCAATC 2147
QY 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856
Db 2148 GCTACAGCTACGGCAGGATTAATTGTGCCAACTCCATTTGTGAGCGATAGAGCAAAAT 2207
QY 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2208 ACGTCCGAACTCTTATTGTTGATCGCGCAAGTGTGGAATACTAATGCAAAATCA 2267
QY 870 -----AspLysGlyPheThrAsn 875
Db 2268 GATAAAATGGATTAGAGCGCATGTATTAAAAAGATTGCTGATTATGCAAAATCAAGC 2327
QY 876 AspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArg 895
Db 2328 CGTATTCCGCCCTCTACAGGTCGGATTCATGGAATGGAATCTCTATTGGGCCATTGGTA 2387
QY 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
Db 2388 TTCTCTTATCCCAACCAATTAATAAATATGAAAAATGATCATGTCGAACACAGTTCCAATTT 2447
QY 914 PheIleGlyThrProPhe 919
Db 2448 AGTATTGAGCGTTCTTTC 2465

```

RESULT 9

US-08-942-046-1

Sequence 1, Application US/08942046

Patent No. 6264954

GENERAL INFORMATION:

APPLICANT: CHONG, Pele

APPLICANT: THOMAS, Wayne

APPLICANT: YANG, Yan Ping

APPLICANT: LOOMORE, Sheena

APPLICANT: SIA, Dwo Yuan Charles

APPLICANT: KLEIN, Michel

TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN

NUMBER OF SEQUENCES: 55

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney

STREET: 6TH Floor, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/942,046
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/433,522
FILING DATE: 12-SEP-1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: STEWART, Michael I.

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-732 MIS:jb

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2949 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 75..2465

US-08-942-046-1

Alignment Scores:
Pred. No.: 2,03e-08 Length: 2949
Score: 172.00 Matches: 171
Percent Similarity: 35.21% Conservative: 141
Best Local Similarity: 19.30% Mismatches: 344
Query Match: 3.64% Indels: 230
DB: 4 Gaps: 44

US-09-914-168-2 (1-919) x US-08-942-046-1 (1-2949)

QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
DB 63 AGGATACATCGATGAAAAAAGTCTTAATCGCAAGTTTATTTCGTACGACAGCACT 122
QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLysGluAsp 218
DB 123 GTGTTGCGCA-----CCTTTGTGGCAAAAGATATTCTGTGTGGATGT 167
QY 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
DB 168 GTTCAGGTGCTTGAACAACAATCCGAGCAAGTTTACCT----- 209
QY 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
DB 210 GTTCGCGCGTCAGCGTGTGACTGACATGATGTG--GCTAATATTGTCGCTCTTTA 266
QY 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
DB 267 TTCGTAAGTGGTTCGATTCGATGTGAAGGCGCATCAAGAAGCGCATGCTGTGTGT 326
QY 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
DB 327 AGCGTTGGTGGTAAATCGATCATTTTCAGATGTAAATCAAAAGTTAATCTCTTATTC 386
QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
DB 387 ACTGAAGCATTAAACAAACTTACATGCTAACGGGTTTAAAGTTGGCGATGTTTAA--- 443
QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
DB 444 -----ATTCGAGAAAAATTAATGAATTTGCCAAAAGTGTAAGAGCACTATGCA 494
QY 329 PheAspGlyArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsn 345

DB 495 AGTGTAGTCGCTAT---AACGCAACAGTTGAACCTATTGTCAATAGCTTACCAATAAT 551
QY 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
DB 552 CGGCTGAAATTTTAATCAATCAATGAAGATGATACCAAAATTTGGCATCACTAACT 611
QY 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
DB 612 TTC-----AAGGGGAACGAA---TCTGTTAGTACGAGTACATTACAAGAA 653
QY 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402
DB 654 CAAATGGAATTACAACCTGATTCTTGGTGGAAATTTATGGGGAATAATAATTTGAAGTGC 713
QY 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
DB 714 CAATTCGAGAAAGATTTCAGTCAATCGT-----GATTAT 749
QY 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
DB 750 TATTTAAATTAATGGCTATGCCAAGCACAATTTACTAAAACGGATGTTCACTAAATGAT 809
QY 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456
DB 810 GAAAAACAAAGTTAATGTAACCATTTGATGTAATGAAGGTTTACAGTATGACCTTCGT 869
QY 457 -----GluSerThrLeuGluProValIleGlu 465
DB 870 AGTGCACGCATTATAGTAATCTGGGAGGTATGCTGCCGAGCTTGAACCTTTACTTTCA 929
QY 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484
DB 930 GCATTACATTTAAATGATATCTCCGCGGTAGTGATATTGCAGATGTAGAA----- 980
QY 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
DB 981 ---AATGCAATTAAAGCAAACTTGGAGACGCGTTACCGTACGCGCAACGGTAAATTC 1037
QY 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
DB 1038 GTACCTGATTTTCATGATGCAATAAACAATAGCGATAACCTTGTGTTGATGCTGGA 1097
QY 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
DB 1098 CGACGTTTAACTGTCGCCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTGATGACT 1157
QY 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
DB 1158 TTACGTCAGGAAATGCCCAACAAGAGAACTTGGTATATTCACAAATAGTTGAGTTA 1217
QY 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
DB 1218 GGAATAATTCGCTTAGATCGTACAGGTTCTTCCGAAACAGTCGAAACCGAATGATCCT 1277
QY 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
DB 1278 ATCAATGGTAGTAATGATGAAGTGCATGCTGATATATAAGTCAAA----- 1322
QY 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592
DB 1323 -----GAACGTACACGCGTAGTATCAACTTTGTTATGTTACCGGT 1364
QY 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608
DB 1365 ACAGAGAGTGTATTAGTTATCAAGCAAGTGTAAACAACAGATAATTTCTGGGACAGGG 1424
QY 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
DB 1425 GCGGAGTAGTATAGCTGGTACGAAAAATGATTATGTCAGGAGTGTCAATTTGCGTTAT 1484
QY 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
DB 638 -----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638


```
Db 1485 ACCGAGCCCTATTATTACTAAAGATGGTGTAAAGTCTTGTCGAAATGTTTCTTTGAAAC 1544
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658
Db 1545 TACGATAACTCTAAAGATGATACATCTCTAACTATATAGCGTAGGACTTACGGAAGTAAT 1604
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671
Db 1605 GTTACTTTAGGTTTCCCTGTAAATGAAATAAATCTCTATTATGTAGGATTAGGTATACC 1664
Qy 672 -----GluLeuSerArgSerIleGlnAsnGlyClyTrpAsnArgThr----- 686
Db 1665 TATAATAAAATAGTAACTTGTCTAGAA-----TATAACCGTAAATTTATATATT 1715
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGln 706
Db 1716 CAATCAATGAAATTAAGGTATGATGCAATTAACAAAT----- 1754
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGlyVal 726
Db 1755 -----GACTTTGATTTTCTTTTGGTTGG----- 1778
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
Db 1779 -----AACTATAACAGCCTTAATAGAGGCTATTTCCTCAACT 1814
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 1815 AAAGGGGTTAAAGCAAGTCTTGGTGACGAGTATCTATTCAGGTTCTGATAACAAATAC 1874
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 1875 TACAACTAAGTGCAGATGATACAGGGTTCTACCCATTA----- 1913
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 1914 GACAGATACACTCTGGTGTGTATCTGCAGAAAGCATCTGCAGGATATGCAATGGTTT 1973
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820
Db 1974 GGAACAACAGCGTTTACCGTTCTATCAAACTTATACAGCGGGTGCATCGGTTCAATAGT 2033
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2034 GCTTTGCTTATGTAGTATTTGGACCTAACGCAATTTATGCCGAATATGTTAATGGTAGT 2093
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842
Db 2094 GGTACTGGTACTTTTAAAGAGATAAGTCTGAT-----GTGATTGGTGGTAATGCAATC 2147
Qy 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856
Db 2148 CCTACAGTACGCGAGAGTTAATTTGTGCCAACTCCATTTGTGAGCGATAAGAGCCAAAT 2207
Qy 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2208 ACGGTCCGAACCTCTTATTTGTGTGTCGCGCAAGTGTTTGGAATATAAATGAAATCA 2267
Qy 870 -----AspLysGlyPheThrAsn 875
Db 2268 GATAAAATGGATPAGAGCGCATGTATTAAAAAGATTGCTGATTATGCAAAATCAAGC 2327
Qy 876 AspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArg 895
Db 2328 CGTATTCGCCCTCTACAGGTGCGGATTCCTCAATGGCAATCTCTATTGGGCATTTGGTA 2387
Qy 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
Db 2388 TTCTCTATATCCCAACCAATTAATAAATATGAAATATGATGATCTCGAACAGTTCCTCAATTT 2447
Qy 914 PheIleGlyThrProPhe 919
Db 2448 AGTATTGGAGTCTTTC 2465
```

```
RESULT 10
US-08-433-522A-5
: Sequence 5, Application US/08433522A
: Patent No. 6013514
: GENERAL INFORMATION:
: APPLICANT: CHONG, Pele
: APPLICANT: THOMAS, Wayne
: APPLICANT: YANG, Yan Ping
: APPLICANT: LOOSMORE, Sheena
: APPLICANT: SIA, Dwo Yuan Charles
: APPLICANT: KLEIN, Michel
: TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN ;
: NUMBER OF SEQUENCES: 55
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: 6TH Floor, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1K7
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA: US/08/433,522A
: FILING DATE: 12-SEP-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: STEWART, Michael I
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-434 MLS:jb
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 5:
: LENGTH: 2950 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 334..2724
US-08-433-522A-5
```

```
Alignment Scores:
Pred. No.: 2,03e-08 Length: 2950
Score: 172.00 Matches: 171
Percent Similarity: 35.21% Conservative: 141
Best Local Similarity: 19.30% Mismatches: 344
Query Match: 3.64% Indels: 230
DB: 3 Gaps: 44
```

```
US-09-914-168-2 (1-919) x US-08-433-522A-5 (1-2950)
Qy 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyClyThrSerAla 198
Db 322 AGGATACATCGATCAAAAACTTCTAATCCAGCTTTATTATTCGGTACGACAAACGACT 381
Qy 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLeuGluAsp 218
Db 382 GTGTTTGGCGCA-----CCTTTGTGCAAAAGATATTCGTGTGCGTGTG 426
Qy 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 427 GTTCAAGGTGACTTAGACAACAATTCGACGACGATTTACCT----- 468
Qy 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 468
```

Db 469 GTTCGTGCGGGTCAACGCTGTGACTGACAATGATGTG---GCTAATAATTGTCGCCGCTCTTTA 525
Qy 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 526 TTCGTAAGTGGTTCGATCGATGATGGAAGCGCATCAAGAAGCGGATGCTGTGTGTT 585
Qy 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
Db 586 AGCGTTGTGGCTAAATCGATCATTTTCAGATGTTAAATCAAAGTAACCTCTGTATTTC 645
Qy 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 646 ACTGAAGCACTAAACAAACTTAGACTAAGCGGTTTAAAGTTGCGCATGTTTAA--- 702
Qy 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 703 -----ATTCGAGAAAAATTAATGAATTTGCCAAAAGTGTAAAGCAGCACTATGCA 753
Qy 329 PheAspGlyArgTrpLeuAspArgSerValaspValIle-----LeuProAspAsn 345
Db 754 ACTGTAGGTTCGGTAT---AACGCAACAGTTGAACCTATTGTCAATACGCTACCAATAAT 810
Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 811 CCGCTGAATTTTAATTCAAATCAATGAAGATGATAAAGCAAAATTTGGCATCATTAAC 870
Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
Db 871 TTC-----AAGGGGAACGAA---TCTGTTAGTAGCAGTACATTACAAGAA 912
Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402
Db 913 CAATGGAATTAACAACCTGATTCCTGGTGGAATATATGGGAAATAAATTTGAAGGTGCG 972
Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 973 CAATTCGAGAAAGATTTGCAGTCAATTCGT-----GATTAT 1008
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 1009 TATTTAAATTAATGGCTATCCCAACGACACAAATTAATAACGCGATGTTTCAGCTAATGAT 1068
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456
Db 1069 GAAAAACAAAGTTAATGAACCTGATGTAATGAATGAAGTTTACAGTATGACCTTCGT 1128
Qy 457 -----GluSerThrLeuGluProValIleGlu 465
Db 1129 AGTGCACGCATTAATAGGTAACTGGGAGGTATGCTGCGCAGCTTGAACCTTTACTTTCA 1188
Qy 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484
Db 1189 GCATTACATTTAAATGATACTTTCGCCCGCTAGTGATATGTCAGATGTAGAA----- 1239
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
Db 1240 ---AATGCAATTAAGCAAAACTTGGAGAACGCGGTTACGGTAGCGCAACGCGTAAATTCA 1296
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
Db 1297 GTACCTGATTTGATGATGCAATAAACAATAGCCGATACCCCTGTTGTTGATGCTGGA 1356
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
Db 1357 CGACGTTTAACTGTTCCGCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTGATAGCACT 1416
Qy 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
Db 1417 TTACGTCAGGAAATCGCCAAACGAAGAACTTGGTATATAATTCAACAATTAGTTAGTTA 1476
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
Db 1477 GGAATAATTCGCTTAGCTGATAGGTTTCTTCGAAACAGTCGAAACCGAATTTGATCCT 1536

Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
Db 1537 ATCAATGGTAGTAATGATGAAGTGGATGCTGATATATAAAGTCAA----- 1581
Qy 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTyrPgly 592
Db 1582 -----GAACGTAAACACGCGTAGTATCAACTTTGGTATTGGTTACGGT 1623
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608
Db 1624 ACAGAGAGGTATTAGTTATCAAGCAAGTGTAAACCAAGATATTTCTTGGGAACAGGG 1683
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 1684 CGGCAGTAGTATAGTCTGCTACGAAAAATGATTATGGTACGAGTGCAATTTGGGTAT 1743
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
Db 1744 ACGAGCCCTATTATTACTAAAGATGGTGAAGTCTTGGTGAATGTTTCTTTTGAANAAC 1803
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658
Db 1804 TACGATAACTCTAAAAGTATACATCCTCTAACTATAAAGCTACGACTTACGGAAGTAAT 1863
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671
Db 1864 GTTACTTTAGGTTTCCTCTGTAATGAAAAAATACTCTATTATGTAGGATTAGGTCATACC 1923
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThr----- 686
Db 1924 TATAATAAAATTAAGTAACTTTGCTCTAGAA-----TATAACCGTAATTTATATATT 1974
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGln 706
Db 1975 CAATCAATGAATTTAAAGTAATGGCATTAACAACAAAT----- 2013
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
Db 2014 -----GACTTTCATTTTCTTTTGGTGG----- 2037
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
Db 2038 -----AATATAACAGCCTTAATAGAGGCTATTTCCCAACT 2073
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 2074 AAGGGGTTAAACCAAGTCTTGGTGACGAGTTACTATTCCAGGTTCTGTATAACAATAAC 2133
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 2134 TACAAACTAAGTCAGATGTACAGGGTTTCTACCCATTA----- 2172
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 2173 GACAGAGATCACTCTCGGTTGTATCTGCAAAAGCATCTCGCAGGATATGCAAAATGGTTT 2232
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyAspGlnSerIleArg 820
Db 2233 GGAACAACGGTTTACCGTTTCTATCAAACTTATACACGGGTGGCATCGGTTCAATAGT 2292
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2293 GGTTTTCCTATGTTAGTATTTGGACCTAACGCAATTTATGCCGAATATGTTAATGGTAGT 2352
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842
Db 2353 GGTACTGGTACTTTAAGAAGATAAGTCTTGAT-----GTGATTTGGTGTAAATGCAATC 2406
Qy 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856
Db 2407 GCTACAGTAGCCGAGGTAAATTGTGCCAACTCCATTTGTGAGCGGATAAGAGCAAAAT 2466

```

QY 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2467 ACGGTCCGAACCTCTTATTGTGATCGCGCAAGTGTGGAATACTAAATGGAATCA 2526
QY 870 -----AspLysGlyPheThrAsn 875
Db 2527 GATAAAATGGATTAGACGGATGATTAAAGATTGGCTGATTATGCAATCAAGC 2586
QY 876 AspThrLysIleGlyAlaGlyValArgTrpAlaSerProValGlyGlnValArg 895
Db 2587 CGTATTCGCCGCTCTACAGGTGCGGATTCGAATGCAATCTCTATTGGCCATGGTA 2646
QY 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
Db 2647 TTCTCTATGCCAAACCAATTAATAATATGAATGATGATGCAACAGTTCCAATTT 2706
QY 914 PheIleGlyThrProPhe 919
Db 2707 AGTATTGGAGGTTCTTTC 2724

RESULT 11
US-09-135-166-5
; Sequence 5, Application US/09135166
; Patent No. 6083743
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: SIA, Dwo Yuan Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6TH Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/135,166
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/433,522
; FILING DATE: 12-SEP-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-829 MTS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2950 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 334..2724
US-09-135-166-5

```

Alignment Scores:

```

Pred. No.: 2,03e-08 Length: 2950
Score: 172.00 Matches: 171
Percent Similarity: 35.21% Conservative: 141
Best Local Similarity: 19.30% Mismatches: 344
Query Match: 3.64% Indels: 230
DB: Gaps: 44

US-09-914-168-2 (1-919) x US-09-135-166-5 (1-2950)
QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 322 AGGATACAATCGATGAAACAACTTCTAATCGCAAGTTTATTATTCGGTACGACACAGCT 381
QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp 218
Db 382 GTGTTTGGCGCA-----CCTTTTGTGGCAAAAGATATTGCTGTGGATGCT 426
QY 219 IleThrGlnGlnSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 427 GTTCAAGGTGACTTAGAACACAAATCCGACCAAGTTTACCT----- 468
QY 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 469 GTTCGTGCGCGTCACCGTGTGACTGACAATGATGTG---GCTAATATTGTCGCTCTTTA 525
QY 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 526 TTCGTAAAGTGTGCGATTCGATGCTGAAGCGCATCAAGAGCGGATGTGCTTTGTT 585
QY 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
Db 586 ACGGTGTGGCTAAATCGATCATTTTCAGATGTTAAATCAAAAGGTAACCTGTTATTCCC 645
QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 646 ACTGAAGCAGCTTAAACAAACTTAGCTGCTAACGGGTTTAAAGTTGGCGATGTTTAA 702
QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 703 -----ATTGCGAAAAAATTAAATCAATTTGCCAAAAGTGTAAAGAGCATATGCA 753
QY 329 PheAspGlyArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsn 345
Db 754 AGTGTAGTCGCTAT---AACGCAACAGTGTGAACCTATTCTCAATAGCTACCAATAAT 810
QY 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 811 CGCGCTGAAATTTTAAATCAATCAAGATCATAAACCAAAATTTGGCATCATTAAC 870
QY 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
Db 871 TTC-----AAGGGGAGCGAA---TCTGTTAGTAGCAGCTACATTACAAGAA 912
QY 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402
Db 913 CAATGCAATTACAACCTGATTCTTGGTGAAATTTATGGGAAATTAATTTGAAGTGCG 972
QY 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 973 CAATTCGAGAAAGATTTCAGTCAATTCGT-----GATTAT 1008
QY 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 1009 TATTTAAATAATGGCTATGCCAAAGCACAAATTTACTTAAACCGGATGTTTCAGCTAATAT 1068
QY 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456
Db 1069 GAAAAACAAAGTTAATGTAACCATTTGATTAATGAAGGTTTACAGTATGACCTTCT 1128
QY 457 -----GluSerThrLeuGluProValIleGlu 465
Db 1129 ACTGCACGCAATTATAGGTAATCTGGGAGGTATGTCTGCGGAGCTTGAACCTTTACTTCA 1188

```

Qy 466 ThrValGluLeuThrAspGlyLeuLeuMetAspIleSerProIleGluPheSerAla 484
Db 1189 GCATTACATTTAAATGATATCTCGCGCTAGTATGATTTGCAGATGTAGAA 1239
Qy 485 SerAsnLeuIleGlnAspLysLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
Db 1240 ---AATGCAATTAAGCAAACTTGGAGAACGCGTTTACGGTAGCGCAACGGTAATAATCA 1296
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
Db 1297 GTACCTGATTTGATGATGCAATAAATACATTTAGCGATAACCTTGTGTTGATCCTGGA 1356
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
Db 1357 CGAGCTTTAACTGTTGCGCAACTTCGCTTGAAGGAATACCGCTTCTGCTGATAGCACT 1416
Qy 534 ValAlaArgAlaIleLeuProAspLysSer-----GluAsnGluValIleAspLeu 550
Db 1417 TTACGTCAGGAATCGGCCAACAGAGAACTTGGTATAATTCACAATTAGTTGAGTTA 1476
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
Db 1477 GGAAAAATTCGCTTAGATCGTCAGGTTTCTTCGAAACAGTCGAAACCGAATTCATCCT 1536
Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
Db 1537 ATCAATGCTAGTAAATGAGTGGATGCTGATATAATAAGTCAAA----- 1581
Qy 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTyrGly 592
Db 1582 -----GAACGTAACACGGGTAGTATCAACTTTGGTATTGGTTACCGT 1623
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe-----GluHisAsnLeuIle 608
Db 1624 ACAGAGAGTGCTATTAGTATCATCAGCAAGTGTAAACAGAGTAATTTCTTGGACACAGG 1683
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 1684 CGGCAGCTAAGTATAGCTGGTACGAAAAATGATTGTAGCAGTGTCAATTTGGGTTAT 1743
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
Db 1744 ACCGAGCCTATTTTACTAAGATGGTGTAGTCTTGTGGAAATGTTTCTTTGAAAC 1803
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPheGlyHisSer 658
Db 1804 TACGATAACTCTAAAAGTAGATACCTCTAACTATAAGCGTACGACTTACGGAAGTAA 1863
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671
Db 1864 GTTACTTTAGGCTTCCCTGTAAATGAAATAACTCCTATTATGTAGGATTAGGTCATACC 1923
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThr----- 686
Db 1924 TATAATAAATTTAGTAACCTTTGCTCTAGAA-----TATAACCGTAATTTATATATT 1974
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLysThrGlnAlaProProGluThrTrpGln 706
Db 1975 CAATCAATGAAATTTAAAGGTAATGGCATTAAACAAT----- 2013
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
Db 2014 -----GACTTTTGGATTTTCTTTTGGTTGG----- 2037
Qy 727 AlaValHisLysThrValAlaAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
Db 2038 -----AATATAACAGCCTTAATAGAGGCTATTTCCTCCAACT 2073
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 2074 AAAGGGTTAAAGCAAGTCTTGGTGGAGGAGTACTATTCCAGGTTCTGTATAACAAATAC 2133
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782

Db 2134 TACAAACTAAGTGCAGATGTACAGGTTTCTTACCCATTA----- 2172
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 2173 GACAGAGATCACTCTCGGTTGTATCTCAAAAGCATCTGCAGGATATGCAAAATGGTTTT 2232
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820
Db 2233 GGAACAACCGTTTACCGTTCTATCAAACTTATACAGCGGGTGGCATCGGTTTACATCGT 2292
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2293 GCTTTTGTCTTATGAGTAGTATGACCTTAACGCAATTTATGCCGAATATGCTAATGGTAGT 2352
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842
Db 2353 GGTACTGGTACCTTTTAAAGAAAGATAGTCTGTAT-----GTGATTGGTGGTAAATGCAATC 2406
Qy 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856
Db 2407 GCTACAGCTAGCGCAGAGTTAATTGTGCCCACTCCATTTGTGAGCGATTAAGACCCAAAT 2466
Qy 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2467 ACGGTCGGAACCTCTTATTTCTGTATGCGCAAGTGTGGAAATCTAAATGGAATCA 2526
Qy 870 -----AspLysGlyPheThrAsn 875
Db 2527 GATAAAATGGATTAGAGCGCATGATTAAAGAAATGCTCCTGATTTATGCAAAATCAAGC 2586
Qy 876 AspThrLysIleGlyAlaGlyValGlyValArgTyrAlaSerProValGlyGlnValArg 895
Db 2587 CGTATTCGCGCCTCTACAGGTGCGGATTCGAATGCAATCTCCTATTGGGCGATTGGTA 2646
Qy 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
Db 2647 TTCTCTTATGCCAACCAATTAATAAATATGAAATGATGATGTCGAACAGTTCCCAATTT 2706
Qy 914 PheIleGlyThrProPhe 919
Db 2707 AGTATTGGAGGTTCTTTC 2724
RESULT 12
US-08-942-046-5
Sequence 5, Application US/08942046
Patent No. 6264954
GENERAL INFORMATION:
APPLICANT: CHONG, Pele
APPLICANT: THOMAS, Wayne
APPLICANT: YANG, Yan Ping
APPLICANT: LOOSMORE, Sheena
APPLICANT: STA, Dwo Yuan Charles
APPLICANT: KLEIN, Michel
TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
NUMBER OF SEQUENCES: 55
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6TH Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/942,046
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: 08/433,522
FILING DATE: 12-SEP-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I.
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-732 MIS:jfb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2950 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 334..2724
US-08-942-046-5

```

Alignment Scores:

```

Pred. No.: 2,03e-08 Length: 2950
Score: 172.00 Matches: 171
Percent Similarity: 35.21% Conservative: 141
Best Local Similarity: 19.30% Mismatches: 344
Query Match: 3.64% Indels: 230
DB: 4 Gaps: 44

```

US-09-914-168-2 (1-919) x US-08-942-046-5 (1-2950)

```

QY 179 LysValProArgLeuLysAlaTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 322 AGGATCAATCGATGAAACAACTCTAATCGCAAGTTTATTTCGGTAGCAACAGACT 381
QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp 218
Db 382 GTGTTTGGCGCA-----CCTTTGTGGCAAAAGATATTCGTGTGGATGGT 426
QY 219 IleThrGlnGlnSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 427 GTTCAAGTGGACTTAGAACCAACAAATCCGACGACGTTTACCT----- 468
QY 239 LeuValAlaAlaAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 469 GTTCGGCGCGTACGGTGTACTGACAAATGATGTG---GCTAATATTGTCCGCTCTTA 525
QY 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 526 TTCGTAAGTGTGCGATTGCGATGATGTGAAGCGCATCAAGNAGCGATGCTGTGTGTT 585
QY 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
Db 586 AGCGTTGTGGCTAATCGATCATTTTCAGATGTTTAAATCAAGTAACGTAACCTGTTATCC 645
QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 646 ACTGAACGACTTAAACAAACTTAGATGCTAACGGGTTTAAAGTTGCCGATGTTTA--- 702
QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 703 -----ATTCCAGAAAATTAATGAATTTGCCAAAAGTGTAAGACGACCATGCA 753
QY 329 PheAspGlyArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsn 345
Db 754 AGTGTAGTGGCTAT---AACGCAACAGTTGAACCTATTGTCAATACGCTACCAATAAT 810
QY 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 811 CGCGCTGAATTTAATCAATCAATGAAGATGAATAAGCAAAATTTGGCATCATTA 870
QY 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385

```

```

Db 871 TTC-----AAGGGGAACGAA---TCTGTTTAGTAGCAGTACATTACAAGAA 912
QY 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402
Db 913 CAAATGAATTAACAACCTGATCTTGGTGGAAATTAATGGGAAATAAATTTGAGGTGCG 972
QY 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 973 CAATTCGAGAAAGATTTTCAGTCAATTCGT-----GATTAT 1008
QY 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 1009 TATTTAAATAATGGCTATGCCAACCAAAATTACTAAACGGATGTTTCAGCTAAATCAT 1068
QY 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456
Db 1069 GAAAAACAAAGTTAATGTAACCATGTAATAAGCTTTTACAGTATGACCTCGT 1128
QY 457 -----GluSerThrLeuGluProValIleGlu 465
Db 1129 AGTGCACGCATTATAGGTAATCTGGGAGGTATCTCCGAGCTTGAACCTTTACTTTCA 1188
QY 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484
Db 1189 GCATTACATTTAATGATGATCTTCCGCCGTAGTATGATATTCAGATCTAGAA----- 1239
QY 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
Db 1240 ---AATGCAATTAAAGCAAACTTGGAGAACGCGGTTACGGTAGCCCAACGGTAAATTTCA 1296
QY 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
Db 1297 GTACCTGATTTTGTATGATGCAATAAAACATTAGCAATAACCTCTGTTGTTGATCTGCA 1356
QY 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
Db 1357 CGACGTTTAACTGTCGCCCAACTTCCTTTGAAGGAATAACCGCTTCTGCTGATGACACT 1416
QY 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
Db 1417 TTACGTCAGAAATGCGCCCAACGAAGGAACCTTGTGATAATTACAAATTAGTTCAGTTA 1476
QY 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
Db 1477 GGAAAAATTCGCTTAGATCGTACAGGTTCTTCGAAACAGTCGAAACCCGAATTGATCCT 1536
QY 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
Db 1537 ATCAATGGTAGTAATGATGAAGTGGATGCTGATATATAAGTCAAA----- 1581
QY 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592
Db 1582 -----GAACGTAACACGGGTAGTATCAACTTTGTTGATTTGTTAGCT 1623
QY 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608
Db 1624 ACAGAGTGGTATTAGTTATCAAGCAAGTGTAAACAAGATAAATTTCTTGGCAACAGCG 1683
QY 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 1684 CGCGCAGTAAGTATAGCTGTACGAAAAATGATTATGTTGAGTGTCAATTTGGCTTAT 1743
QY 622 SerGlu-----AspLysLysGlyValLysLeuThrAlaThrLysProLeuSerHis 638
Db 1744 ACCGAGCCCTATTTTACTTAAAGATGGTGAAGTCTTGGTGGAAATGTTTTCTTCAAAC 1803
QY 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHis 658
Db 1804 TACCATAAATCTCTAAAGTACATCTCTTAATAAGCGTACGACTTACGGAAGTAAT 1863
QY 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671

```


Db 626 AGCGTTGTGGCTAAATCGATCATTTTCAGATGTTAAATCAAAAGGTAACCTCTCTTTATCC 685
Qy 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 686 ACTGAAGCAGCTTAAACAAACCTTGAATGCTAAGCGTTTAAAGTTGGCGATCTTTA --- 742
Qy 309 HisGlyLysTyrGluThrLysLysAlaLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 743 -----ATTTCGAGAAATAATTAATGAATTTGCGAAAGTGTAAAGAGCAGCACTATGCA 793
Qy 329 PheAspGlyArgTyrLeuAspArgSerValAspValIle-----LeuProAspAsn 345
Db 794 AGTGTAGTCGTAT---AACCCACAGTTGAACCTATTGTCATACGCTACCAATAAT 850
Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 851 CGCGCTGAATTTAATCAATCAATCAAGATGATAAGCAAAATTTGCATCATTA 910
Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
Db 911 TTC-----AAGGGGAACGAA---TCTGTGTAGTAGCAGTACATTACAGAA 952
Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402
Db 953 CAATGGAATACAACTGATTCCTGGTGGAAATTTATCGGGAATAAATTTGAAGTCCG 1012
Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 1013 CAATTCGAGAAAGATTTGACGTCAATTCGT-----GAATTAT 1048
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 1049 TATTTAATAATGCTATGCCAAAGCACAAATTTACTAAACGGATGTTTCAGCTAAATGAT 1108
Qy 439 GlnValSerPheGluGlnSerSerArgThrGluProAlaGlnValAsp----- 456
Db 1109 GAAAAACAAAGTAAATGTAACCATGATGTAATGAAGGTTTACAGATGACCTTCGT 1168
Qy 457 -----GluSerThrLeuGluProValIleGlu 465
Db 1169 AGTCAGCATATTAGTAATCTGGGAGGTATGCTCCGAGCTTGAACCTTTACTTTCA 1228
Qy 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484
Db 1229 GCATTACATTTAAATGATACTTCCGCGTAGTGATATGCAGATGTAGAA----- 1279
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
Db 1280 ---AATGCAATTAAGCAAAACTTGGAGAACGCGGTTACGGTACGCGCAACGGTAAATTC 1336
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
Db 1337 GTACCTCATTTGATGATGCAATAAACAATACCGATACCCCTGTGTGTGTGCTGGA 1396
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
Db 1397 CGAGGTTTAACTGTTCCCAACTTCGCTTTGAAGGAATAACCGTTTCTGCTGATAGCACT 1456
Qy 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
Db 1457 TTACGTGAGGAATGCGCCCAACAGAGAACTTGTGTATTAATTCACAATTAGTTGAGTTA 1516
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
Db 1517 GGAATAATTCGCTAGATCGTACAGGTTTCTCGAACAACGCTGAAACCGAATTTGATCCT 1576
Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
Db 1577 ATCAATGGTAGTAATGATGAAGTGGATGTCGTATATAAAGTCAA----- 1621
Qy 575 ValPheValAlaSerAspLysProArgAspGlyClnIle-----GlyLeuGlyTyrGly 592
Db 1622 -----GAACGCTAACGGGGTAGTATCAACTTTTGGTATTGGTTACGGT 1663

Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608
Db 1664 ACAGAGAGTGGTATTAGTTATCAAGCAAGCTGTAAACAGATAAATTTCTGGCAACAGGG 1723
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 1724 GCGGAGTAAGTATAGCTGGTACGAAATGATTATGTGTACGAGTGTCAATTTGGTTAT 1783
Qy 622 SerGlu-----AspLysLysValLysLeuTyrAlaThrLysProLeuSerHis 638
Db 1784 ACCGAGCCCTATTTTACTAAAGATGCTGTAAAGTCTGTGGAAATGTTTCTTCTTCAAAAC 1843
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658
Db 1844 TACGATACTCTAAAGTGATACATCCCTCACTATAAGCGTACGAGTACCGAAGTA 1903
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671
Db 1904 GTTACTTTAGGTTCCCTGTAAATCAAAATCACTCTATTATGTAGGATTAGGTCAATACC 1963
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTyrPheAsnArgThr--- 686
Db 1964 TATAATAAATTAGTAACCTTTGCTCTAGAA-----TATAACCGTAATTTATATATT 2014
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGln 706
Db 2015 CAATCAATGAATTTAAAGGTAATGGCAATTAACAAAT----- 2053
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
Db 2054 -----GACTTTGATTTTTCTTTGGTGG----- 2077
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
Db 2078 -----AACTATAACAGCTTAAATAGAGCTATTTCCCAACT 2113
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 2114 AAAGGGTTAAAGCAAGCTTGTGGTGGAGGATTACTATTCCAGGTTCTGATAACAATAC 2173
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 2174 TACAACTAAGTCAGATGATACAGGTTTCTACCCATTA----- 2212
Qy 783 AsnArgAlaHis-----GlnMetThrGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 2213 GACAGAGATCACCTCTCGGTTGTATCTGCAAAAGCATCTCGAGGATATGCAATGGTTTT 2272
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820
Db 2273 GGAACAAGCGTTTACCGTTCTATCAAACTTATACAGCGGTGCGCATCGTTTCATTACGT 2332
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2333 GGTTCCTTATGCTAGTATTGGACCTACGCCAATTTATGCCGAATATGGTAATGGTAGT 2392
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGluValLeu 842
Db 2393 GGTACTGGTACTTTTAAAGAGATAAGTCTCTGAT-----GTGATTGGTGTATGCAATC 2446
Qy 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp--- 856
Db 2447 GCTACAGCTAGCGCAGAGTTAATTTGTGCCAACTCCATTTGTGAGCGATAAGAGCCAAAT 2506
Qy 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2507 ACGTCCGAACTCCTTATTTGTTGATGCGGCAAGTGTTTTGGAAATACTAAATGCAATCA 2566
Qy 870 -----AspLysGlyPheThrAsn 875
Db 2567 GATAAAATGGATTACAGAGCGGATGTTATTAANAAGATTGCCTGATTATGCAANTCAAGC 2626

Qy 876 AspThrLysIleGlyAlaGlyValGlyValArgTyrPalaSerProValGlyGlnValArg 895
Db 2627 CGTATTGCGGCTCTACAGGTGTCGATTCCAAATGCAATCTCCATTGGGCAATTGGTA 2686
Qy 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
Db 2687 TTCCTTATGCCAAACCAATTAATAAATATGAAATGATGATGTCGAACAGATTCCAATTT 2746

Qy 914 PheIleGlyThrProPhe 919

Db 2747 AGTATTGGAGGTCTCTTC 2764

RESULT 14

US-09-135-166-3

; Sequence 3, Application US/09135166

; Patent No. 6083743

; GENERAL INFORMATION:

; APPLICANT: CHONG, Pele

; APPLICANT: THOMAS, Wayne

; APPLICANT: YANG, Yan Ping

; APPLICANT: LOOSMORE, Sheena

; APPLICANT: SIA, Dwo Yuan Charles

; APPLICANT: KLEIN, Michel

; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN

; NUMBER OF SEQUENCES: 55

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: 6TH Floor, 330 University Avenue

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: M5G 1K7

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA: US/09/135.166

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/433,522

; FILING DATE: 12-SEP-1995

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: STEWART, Michael I

; REGISTRATION NUMBER: 24,973

; REFERENCE/DOCKET NUMBER: 1038-829 MIS:jb

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (416) 595-1155

; TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2984 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 374..2764

; US-09-135-166-3

Alignment Scores:

Pred. No.:	2,07e-08	Length:	2984
Score:	172.00	Matches:	171
Percent Similarity:	35.21%	Conservative:	141
Best Local Similarity:	19.30%	Mismatches:	344
Query Match:	3.64%	Indels:	230
DB:	3	Gaps:	44

US-09-914-168-2 (1-919) x US-09-135-166-3 (1-2984)

Qy 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 362 AGGATACAACTCGATGAAAAAACTTCTAATCGCAAGTTTATTATTCGGTAGACAACGACT 421
Qy 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLeuGluAsp 218
Db 422 GTGTTTGGCGCA-----CCTTTTGGCAAAAGATATTTCGTGGATGGT 466
Qy 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 467 GTTCAAGGTGACTTAGAACACAAATCCGAGCAAGTTTACCT----- 508
Qy 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleLeuArgAsnSer 258
Db 509 GTTCGTGCGCGGTGACGCTGACTGACAAATGATGATG--GCTAATATTGCTCGCTCTTTA 565
Qy 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 566 TTCGTAAGTGGTTCGATTTCGATGATGTAAGCGCATCAAGAGCGGATGTGCTTTGTTGT 625
Qy 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
Db 626 AGCGTTGTCGTAATTCGATCATTTTCAGATGTTTAAATCAAGGTAACCTCTGTATTCCC 685
Qy 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 686 ACTGAAGCACTTAAACAAAACTTAGATGCTAACGGGTTTAAAGTGGCGATGTTTA-- 742
Qy 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 743 -----ATTCGAGAAAAATTAATGAATTTGCCAAAAGTGTAAGAGAGCACTATGCA 793
Qy 329 PheAspGlyArgTyrLeuAspArgSerValAspValIle-----LeuProAspAsn 345
Db 794 AGTGTAGGTGCGTAT--AACGCAACAGTTGAACCTTGTCTCAATACGCTACCAATAAT 850
Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 851 CGCGCTGAAATTTTAATCAATCAATGAAGATGAAGCAAAATTTGGCATCATTAAC 910
Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrValAsnMetGlyGluAlaTyr----- 402
Db 911 TTC-----AAGGGCAACGAA---TCTGTTAGTAGCAGCTAGATTACAAGAA 952
Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402
Db 953 CAATTCGAGAAAGATTTGCAGTCAATTCGT-----GATTAT 1048
Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 1013 CAATTCGAGAAAGATTTGCAGTCAATTCGT-----GATTAT 1048
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 1049 TATTTAAATAATGGCTATGCCAAAGCACAAATTTACTAAAACGAGGATGTTCAGCTAAATGAT 1108
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456
Db 1109 GAAAAACAAAAGTTAATGTAAACCATGTGTAATGAAGGTTTACAGTATGACCTTCGT 1168
Qy 457 -----GluSerThrLeuGluProValIleGlu 465
Db 1169 AGTGCACGCATTATAGTAATCTGGGAGGTATGCTGCCGCGAGCTTGAACCTTTACTTTCA 1228
Qy 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484
Db 1229 GCATTACATTTAAATGATATCTTTCCGCGGTAGTGATATTCAGATGTAGAG----- 1279
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
Db 1280 ---AATGCATTTAAACAAAACCTTGGAGAACGCGGTTACCGTAGCGCAACGGTAATTC 1336

QY 505 MetProasp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
Db 1337 GTACCTGATTGATGCAATAAATACATAGCGATAACCCCTGTTGTTGATGCTGGA 1396
QY 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
Db 1397 CGAGCTTAACTGTTCCCAACTTCGTTTGAAGGAATAACCGTTTCTGCTGATAGCACT 1456
QY 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
Db 1457 TTACGTCAGGAATAGCGCCCAACAAGAAGAACTTGCTATATATTCACAATAGTTGAGTTA 1516
QY 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
Db 1517 GGAATAATTCGCTTAGCTAGCTACAGCTTCTTCGAAACAGTTCGAAACCGAATTCATCCT 1576
QY 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
Db 1577 ATCAATGGTAGTAATGATGAAGTGGATGCTATATAAGTCATAA 1621
QY 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592
Db 1622 -----GAAAGTAAACAGCGGTAGTATCAACTTTGGTATTGTTACGGT 1663
QY 593 SerAspThrGlyThrArgLeuValThrLysPhe-----GluHisAsnLeuIle----- 608
Db 1664 ACAGAGAGTGGTATTAGTTATCAAGCAAGTGTAAACACAGATAATTTCTTGGGAACAGGG 1723
QY 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 1724 GCGGAGTAGTATAGCTGGTACGAAATATGATGTTACGAGTCTCAATTTGGGTTAT 1783
QY 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
Db 1784 ACCGAGCCCTATTTACTAAAGATGGTGAAGTCTTGGTGGAATGTTTCTTTCGAAAC 1843
QY 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGluValPheGlyHisSer 658
Db 1844 TACGATAACTCTAAAGTGCATACCTCTAACTATAAGCTAGCAGTACCGAAGTAAAT 1903
QY 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671
Db 1904 GTTACTTTAGGTTCCCTGTTAAAGTAAATAACTCTTATATGATGAGGATTAGTCAATACC 1963
QY 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTyrPsnArgThr----- 686
Db 1964 TATAATAAATAGTAACTTTGCTCTAGAA-----TATAACCGTAATTTATATATT 2014
QY 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGln 706
Db 2015 CAATCAATGAATTTAAAGGTAATGGCATTAACAAAT----- 2053
QY 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
Db 2054 -----GACTTTGATTTTCTTTGGTTGG----- 2077
QY 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
Db 2078 -----AACTATAACAGCCTTAATAGAGGCTATTTCCCAACT 2113
QY 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 2114 AAAGGGTTAAAGCAAGCTTTGGTGGAGGAGTACTATTCAGGTTCTGTATAACAATAC 2173
QY 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 2174 TACAAACTAAGTCAGATGCTACAGGTTTCTACCCATTA----- 2212
QY 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 2213 GACAGAGATCACCTCTGGGTTGATCTGCAAAAGCATCTGAGGATATGCAATGCTTTT 2272
QY 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820

Db 2273 GGAACAACGCGTTTACCGTTCTATCAACAGTGTATACAGCGGTGGCATCGGTTTCATTACGT 2332
QY 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2333 GGTTTTGCTTATGCTAGTATTGGACCTAACCGAATTTATGCGCAATATGTTAATGGTAGT 2392
QY 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842
Db 2393 GGTACTGGTACTTTTAAAGAAGATAAGTTCTGAT-----GTGATTGGTGGTAATCAATC 2446
QY 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856
Db 2447 GCTACAGTAGCGCAGAGTTAATTTGTGCCAATCCATCTCTATTGGGCCATTGGTA 2506
QY 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2507 ACGGTCCGAACCTCTTATTTGTTGATCGGCAAGTGTTCGAATACTAAATGGAATCA 2566
QY 870 -----AspLysGlyPheThrAsn 875
Db 2567 GATAAAATGATTAGAGAGCGATGTATTAAGAGATTGCTGATTATGGCAATCAAGC 2626
QY 876 AspThrLysIleGlyAlaGlyValArgTrpAlaSerProValGlyGlnValArg 895
Db 2627 CGTATTCGCGCTCTACAGGTGCGGATTCCAATGGCAATCTCTATTGGGCCATTGGTA 2686
QY 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuIleSph 913
Db 2687 TTCTCTTATGTCACCAACCAATTAATAAATATGAAATGATGATGTCGAACAGTTCCTCAATTT 2746
QY 914 PheIleGlyThrProPhe 919
Db 2747 AGTATTGAGGTTCTTTC 2764

RESULT 15

US-08-942-046-3
; Sequence 3, Application US/08942046
; Patent No. 6264954
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOSMORE, Sheena
; APPLICANT: SIA, Dwo Yuan Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6TH Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/942,046
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/433,522
; FILING DATE: 12-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-732 MIS:jb
; TELECOMMUNICATION INFORMATION:


```
Db 2015 CAATCAATGAATTAAGAGTAATGGCATTAAACAAAT----- 2053
QY 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
Db 2054 -----GACITTTGATTTTCTTTGGTGG----- 2077
QY 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGluArg 746
Db 2078 -----AACTATAACAGCCTTAATAGAGGCTATTTCCCAACT 2113
QY 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 2114 AAAGGGTTAAAGCAAGCTTTGGTGGAGGAGTACTATTCAGGTTCTGTATAACAATAC 2173
QY 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 2174 TACAACTAAGTCAGATGTACAGGGTTTCTACCCATTA----- 2212
QY 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 2213 GACAGAGATCACCTCTGGTGTGTATCTGCAAAAGCATCTGCAGGATATGCAAAATGGTTT 2272
QY 801 AsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 820
Db 2273 GGAACAAGCGTTTACCGTCTCTATCAACTATACACGGGTGCATCGGTTCATACGT 2332
QY 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2333 GGTTCCTTATGCTAGTATTGTCGACCTAACGCAATTTATGCGCAATATGGTAATGGTAGT 2392
QY 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842
Db 2393 GGTACTGGTACTTTTAAGAAGATAAGTCTGAT-----GTGATTGGTGTAAATGCAATC 2446
QY 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856
Db 2447 GCTACAGCTAGCGCAGAGTAATTTGTCCCACTCCATTTGTGACGATAGAGCCAAAT 2506
QY 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2507 ACGTCGGAACCTCTTATTATTGTTGATGCGGCAAGTGTGGAATACTAAATGGAATCA 2566
QY 870 -----AspLysGlyPheThrAsn 875
Db 2567 GATAAAATGATTAGAGAGGATGTATTAAGAGATTGCTGATATGGCAAAATCAAGC 2626
QY 876 AspThrLysIleGlyAlaGlyValArgTyrAlaSerProValGlyGlnValArg 895
Db 2627 CGTATTCGCGCTCTACAGGTGTCGGATTCCAATGGCAATCTCTATTTGGCCATTGGTA 2686
QY 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
Db 2687 TTCTCTTATGCCAAACCAATTAATAAATATGAAAAATATGATGTGCGAACAGTTCCAATT 2746
QY 914 PheIleGlyThrProPhe 919
Db 2747 AGTATTGGAGGTTCTTTC 2764
RESULT 16
US-08-433-522A-55
; Sequence 55, Application US/08433522A
; Patent No. 6013514
; GENERAL INFORMATION:
; APPLICANT: CHONG, Pele
; APPLICANT: THOMAS, Wayne
; APPLICANT: YANG, Yan Ping
; APPLICANT: LOOMORE, Sheena
; APPLICANT: SIA, Dwo Yuan Charles
; APPLICANT: KLEIN, Michel
; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
```

```
; STREET: 6TH Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/433,522A
; FILING DATE: 12-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-434 MIS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2987 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-433-522A-55
Alignment Scores:
Pred. No.: 2,08e-08 Length: 2987
Score: 172.00 Matches: 171
Percent Similarity: 35.21% Conservatives: 141
Best Local Similarity: 19.30% Mismatches: 344
Query Match: 3.64% Indels: 230
DB: 3 Gaps: 44
US-09-914-168-2 (1-919) x US-08-433-522A-55 (1-2987)
QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
Db 377 AGGATACATCGATCAAAAACCTTAATCGCAAGTTTATTATTTCGGTACGACACGACT 436
QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLeuGluAsp 218
Db 437 GTGTTTCCCGCA-----CCTTTTCTGGCAAGATATTCTGTGGATCGT 481
QY 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 482 GTTCAAGGTGACTTAGAACACAATCCGACGACAGTTTACCT----- 523
QY 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 524 GTTCGTGCGCGTACAGCGTGTGCTGACTGACAAATGATGTG---GCTAATATTGTCGCTCTTTA 580
QY 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
Db 581 TTCGTAAGTGTGCGATTCGATGTCGAAAGCGCATCAAGAGCGCATGCTGCTTGTGTT 640
QY 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
Db 641 AGCGTTGCGCTAAATTCGATCATTTTCAGATGTTAAATCAAGAGGTAACCTGCTTATTCCC 700
QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
Db 701 ACTGAAGCAGCTTAAACAAACTTAGATCTAACGGGTTTAAAGTTGGCGATGTTTAA--- 757
QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
Db 758 -----ATTCGAGAAAAAATAAATGAATTTGCAAAAGTGTAAAAGAGCAGCATGCA 808
QY 329 PheAspGlyArgTyrProLeuAspArgSerValAspValIle-----LeuProAspAsn 345
```

Db 809 AGTGTAGTTCGCTAT---AACGCAACACAGTTGAACCTATTGTCAATACGCTACCAAAATAAT 865
Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
Db 866 CGCGCTGAATTTAAATCAATCAATCAAGATGATAAGCAAAATTCGCATCATTAAT 925
Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
Db 926 TTC-----AAGGGGAACGAA---TCTGTTAGTAGCAGTACATTACAAGAA 967
Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402
Db 968 CAAATGGAATTTACACCTGATCTTGGTGGAAATTTATGGGGAATAAATTAATGAAGTCGC 1027
Qy 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 1028 CAATTCGAGAAAGATTTCGATGCAATTCGT-----GATTAT 1063
Qy 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
Db 1064 TATTTAAATAATGGCTATGCCAAAGACACAAATTTACTAAACGGATGTTTCAGCTAAATGAT 1123
Qy 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456
Db 1124 GAAAAACAAAAGTTAATGTAACCATGATGATGTAATGAAGGTTTACAGATGACCTTCGT 1183
Qy 457 -----GluSerThrLeuGluProValIleGlu 465
Db 1184 AGTCACCCCATATAGGTAATCTGGAGGTATGTCGCGGAGCTTGAACCTTACTTTCA 1243
Qy 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484
Db 1244 GCATTACATTTAAATGATGACTTTCGCGCGTAGTGATATTGCAGATGTAGAA----- 1294
Qy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
Db 1295 ---AATGCAATTTAAAGCAAACTTCGGAGACGCGTTACGGTAGGCAACCGTAAATTC 1351
Qy 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
Db 1352 GTACCTGATTTTGATGATGCAAAATAAACATTTAGCGATAATTCAGATGTAGAA----- 1411
Qy 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
Db 1412 CGACGTTTAACTGTTTCGCCAACTTCGCTTTGAAGGAAATACCGTTTCTGCTGATACGCT 1471
Qy 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
Db 1472 TTACGTCAGGAATCGGCCCAACAGAGAACTTGGTATATATTCACAAATTAGTTGAGTTA 1531
Qy 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
Db 1532 GGAATAATTCGCTTAGATCGTACAGTTTCTTCGAAACAGTCGAAACCGAATTCATCCT 1591
Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
Db 1592 ATCAATGTTAGTAAATGAAAGTGGATGCTGTATATAAGTCAAA----- 1636
Qy 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592
Db 1637 -----GAACGTAACACGGGTAGTATCAACTTGGTATGTTGGTTACCGT 1678
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle----- 608
Db 1679 ACAGAGAGTGGTATTAGTTATCAACCAAGTGTTAAACAAGATAAATTTCTTGGGAACAGG 1738
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 1739 GCGGCAGTAAGTATAGCTGGTACGAAAAATGATTGTTACGAGTGTCAATTTGGGTTAT 1798
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
Db 1799 ACCGAGCCCTATTATTACTAAAGATGGTGTAGCTTTGGTGGAAATGTTTCTTCTTGGAAAC 1858

Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658
Db 1859 TACGATAACTCTAAAGTCATACATCCTCTAACTATAAGCGTACGACTTACGGAAGTAAT 1918
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis--- 671
Db 1919 GTTACTTTAGGTTTCCCTGTAATGAAATAACTCCTATTATGTAGGATAGGTATCATACC 1978
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThr----- 686
Db 1979 TATAATAAAATAGTAACCTTTCCTCTAGAA-----TATAACCGTAATTTATATATT 2029
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLysLeuLysThrGlnAlaProGluThrTrpGln 706
Db 2030 CAATCAATCAATTTAAAGGTAATGCGATTAAACAAAT----- 2068
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
Db 2069 -----GACTTTGATTTTCTTTTGGTTGG----- 2092
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
Db 2093 -----AACTATAACAGCCTTAATAGAGGCTATTTCCTCCAACT 2128
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 2129 AAAGGGTTAAAGCAAGTCTTGGTGCAGGTTACTATTTCAGGTTCTGATAACAATAAC 2188
Qy 763 AlaIleAlaArgAlaGlyLeSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 2189 TACAACAACTAAAGTCAGATGTACAGGTTTCTACCCATTA----- 2227
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 2228 GACAGAGATCACCTCTGGGTTGTATCTGCAAAAGCATCTGCAGGATATATCCAAATGGTTT 2287
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPheAlaGlyGlyAspGlnSerIleArg 820
Db 2288 GGAACAACAGCGTTTACCGTTCTCTCAAACTTATACAGCGGTGCGTACGTTACCT 2347
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2348 GGTGTTTGTCTTATCGTAGTATTGGACCTAACGCAATTTATGCCCAATATGTTATGGTAGT 2407
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842
Db 2408 GGTACTGTTACTTTTAAAGAGATAAGTTCTGAT-----GTGATGGTGGTAAATGCAATC 2461
Qy 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856
Db 2462 GCTACAGCTAGCGCAGAGTTAATTTGCCCAACTCCATTTGTGAGCGATAAGAGCCCAAAAT 2521
Qy 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2522 ACGGTCGCAACCTCTTATTGTTGATGCGCAAGTGTGGAATACTAAATGGAATCA 2581
Qy 870 -----AspLysGlyPheThrAsn 875
Db 2582 GATAAAATGGATTAGAGAGCGATGATTAAANAAGATTGCTGATTATGCGCAATCAAGC 2641
Qy 876 AspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArg 895
Db 2642 CGTATTCGCGCTCTACAGGTCTCGATGTCGAATGGAATCTCCTATTGGCCATTGGTA 2701
Qy 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
Db 2702 TTCTCTTATGCCAACCAATTAATAAATATGAAATGATGATGTCGAACAGTTTCCAATTT 2761
Qy 914 PheIleGlyThrProPhe 919
Db 2762 AGTATGGAGGTTCTTTC 2779

RESULT 17

US-09-135-166-55
 ; Sequence 55, Application US/09135166
 ; Patent No. 6083743
 ; GENERAL INFORMATION:
 ; APPLICANT: CHONG, Pele
 ; APPLICANT: THOMAS, Wayne
 ; APPLICANT: YANG, Yan Ping
 ; APPLICANT: LOOSMORE, Sheena
 ; APPLICANT: SIA, Dwo Yuan Charles
 ; APPLICANT: KLEIN, Michel
 ; TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
 ; NUMBER OF SEQUENCES: 55
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sim & McBurney
 ; STREET: 6TH Floor, 330 University Avenue
 ; CITY: Toronto
 ; STATE: Ontario
 ; COUNTRY: Canada
 ; ZIP: M5G 1R7
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/135,166
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/433,522
 ; FILING DATE: 12-SEP-1995
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: STEWART, Michael I
 ; REGISTRATION NUMBER: 24,973
 ; REFERENCE/DOCKET NUMBER: 1038-829 MIS:jb
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (416) 595-1155
 ; TELEFAX: (416) 595-1163
 ; INFORMATION FOR SEQ ID NO: 55:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2987 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear

US-09-135-166-55

Alignment Scores:

Pred. No.:	2.08e-08	Length:	2987
Score:	172.00	Matches:	171
Percent Similarity:	35.21%	Conservative:	141
Best Local Similarity:	19.30%	Mismatches:	344
Query Match:	3.64%	Indels:	230
DB:	3	Gaps:	44

US-09-914-168-2 (1-919) x US-09-135-166-55 (1-2987)

QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerGlnSerGlyGluThrSerAla 198
 Db 377 AGGATACAATCGATGAAAAAATCTTAATCGCAAGTTTATTATTCGGTAGCAGCAACT 436
 QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaLysGluAsp 218
 Db 437 GTGTTGCCCA-----CCTTTGTGGCAAAAGATATTCTGTGGATGCT 481
 QY 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
 Db 482 GTTCAAGGTGACTTAGAACCAAAATCCGCGCAAGTTTACCT----- 523
 QY 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleArgAsnSer 258
 Db 524 GTTCGTCCCGTCAGCGGTGACTGACAATGATGTG--GCTAATATTCTCGCTCTTTA 580

QY 259 Ile-----GlyGluVal---AspValIleIleHisAspLeuGlyGluProValTyrIle 275
 Db 581 TTCGTAAGTGGTTCGATTCGATGATGTAAGCGCATCAAGAGCGGATGCTGTGTTGTT 640
 QY 276 AspTyrArgAla-----ValGluValArgGlyGluGlyAla---Asp 288
 Db 641 ACGGTTGTCGCTAAATCGATCATTTTCAGATGTTAAATCAAAGGTAACCTGTTATTCCC 700
 QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
 Db 701 ACTGAAGCACTTAAACAAACTTAGATCTAACGGCTTTAAAGTTGGGATGTTTAA--- 757
 QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
 Db 758 -----ATTCGAGAAAAATTAATCAATTTGCCAAAAGTGTAAAGAGCACATATGCA 808
 QY 329 PheAspGlyArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsn 345
 Db 809 AGTGTAGTCGCTAT---AACGCAACAGTTGAACCTATTGTCAATACGCTACCAAAATAAT 865
 QY 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
 Db 866 CGCCTCAAAATTTAATTCAAATCAATCAAGATGATAAGCAAAATTTGGCATCATTAAC 925
 QY 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
 Db 926 TTC-----AAGGGGACGAA---TCTCTTAGTAGCAGTACATTTACAAGAA 967
 QY 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr----- 402
 Db 968 CAATGGAATTACAACCTGATTCTTGGTGCAAAATTTATGGGAAATTAATTTGAAGGTGCG 1027
 QY 403 -----AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
 Db 1028 CAATTCGAGAAAGATTTCAGTCAATTCGT-----GATTAT 1063
 QY 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
 Db 1064 TATTTAAATTAATGGCTATGCCAAAGCACAAATTTACTAAAACGGATGTTTCAGCTAATGAT 1123
 QY 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp----- 456
 Db 1124 GAAAAACAAAGTTAATGTAACCATTTGATGTAATGAAGTTTACAGTATGACCTTCCT 1183
 QY 457 -----GlusterThrLeuCluProValIleGlu 465
 Db 1184 AGTGCACGCATTATAGGTAATCTGGGAGGTATGTCTGCCGAGCTTCAACCTTTACTTTCA 1243
 QY 466 ThrValGluLeuThrAspGlyIle---LeuMetAspIleSerProIleGluPheSerAla 484
 Db 1244 GCATTACATTTAAATGATACTTTCCGCCCTAGTATGATTTGCAGATGTAGAA----- 1294
 QY 485 SerAsnLeuIleGlnAspLysLysLeuValAlaLysAlaLysAlaArgHisLeuTyrAsp 504
 Db 1295 ---AATGCAATTAAGCAAAACTTCGAGAACCGGTTACGGTAGCGCAACGCAAAATTTCA 1351
 QY 505 MetProAsp-----AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
 Db 1352 GTACCTGATTTTGTATGATGCAAAATAAAACATTAGCGATACCCCTTGTGTTGATCCTCGA 1411
 QY 521 ArgSerIleLeu-----GlyArgIleSerAspAlaValSerAla 533
 Db 1412 CGACCTTTAACTGTTCGCCAACTTCGCTTTGAAGGAATAACCGTTTCTGCTGTAGCACT 1471
 QY 534 ValAlaArgAlaIleLeuProAspGluSer-----GluAsnGluValIleAspLeu 550
 Db 1472 TTACGTCAGGAAATCGGCCAACCAAGAAAGAACTTGGTATATAATTCACAAATTAGTTGAGTTA 1531
 QY 551 Pro-----GluArgThr-----AlaLeuAlaAsnArgLysThrPro 562
 Db 1532 GAAAAAATTCGCTTAGATCGTACAGGTTTCTTCGAAAACAGTCGAAAAACCGAATTTGATCCT 1591

Qy 563 AlaAsp-----ValTyrGlnSerLysLysValProLeuTyr 574
Db 1592 ATCAATGGTAGTAATGATGAAGTGGATCGTATATAAAGTCAAA-----1636
Qy 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle-----GlyLeuGlyTrpGly 592
Db 1637 -----GAAACGTAACACGGGTAGTAGTCAACTTTGGTATTGGTTACGGT 1678
Qy 593 SerAspThrGlyThrArgLeuValThrLysPhe---GluHisAsnLeuIle-----608
Db 1679 ACAGAGAGTGGTATTAGTATCAAGCAAGTGTAAACAAAGATAATTTCTGGGAACAGGG 1738
Qy 609 -----AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 1739 GCGGCAGTAAGTATAGTGGTACGAAATAATGATTGTAGGAGTGTCAATTTGGGTTAT 1798
Qy 622 SerGlu-----AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
Db 1799 ACCGAGCCCTATTTTACTAAAGATGGGTGAAGTCTTGGTGGAAATGTTTCTTTGAAAC 1858
Qy 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPheGlyHisSer 658
Db 1859 TACGATRACTAAAGGTATACATCTTAAGTATAGCGTACGACTTACGGAAGTAAAT 1918
Qy 659 ---ThrAsnGlyPheAspLeuSerThrArgThr-----LeuGluHis---671
Db 1919 GTTACTTTAGGTTTCCCTGTAATGAAATAACTCCTATTATGTAGGATTAGGTCATACC 1978
Qy 672 -----GluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThr-----686
Db 1979 TATAATAAATAGTAACCTTGGCTCTAGAA-----TATAACCGTAATTTATATATT 2029
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTyrGln 706
Db 2030 CAATCAATGAATTTAAGGTAAATGGCATTAACAAAT-----2068
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGlnAlaLeuLeuAlaGlyVal 726
Db 2069 -----GACTTTGATTTTCTTTGGTGG-----2092
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
Db 2093 -----AACTATAACAGCCTTAATAGAGGCTATTATAGGCTATTTCCCACT 2128
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsnMet 762
Db 2129 AAAGGGTTAAAGCAAGTCTTGGTGGACGAGTTACTATTCCAGGTTCTGATAACAAATAC 2188
Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 2189 TACAACTAAGTGCAGATGTACAGGGTTTCTACCCATTA-----2227
Qy 783 AsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800
Db 2228 GACAGAGATCACTCTGGTGTATCTGCAAAAGCATCTGCAGGATATGCAATGGTTTT 2287
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPheAlaGlyGlyAspGlnSerIleArg 820
Db 2288 GGAACAACAGCGTTTACCCTTCTATCAAACTATACAGCGGGTGGCATCGGTTTCATTACGT 2347
Qy 821 GlyTyrAlaHisAspSerLeuSerPro-----829
Db 2348 GGTTCCTTATGTAGTATTGGACCTAACGCAATTTATGCCGAATATGGTAATGGTAGT 2407
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu 842
Db 2408 GGTACTGTACTTTTAAAGAGATAGTTCTGAT-----GTGATGGTGGTAATCCAATC 2461
Qy 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp-----856
Db 2462 GCTACAGTAGCGCAGAGTAAATCTGTGCCAACTCCATCTTGTGAGGATAGAGCCAAAT 2521
Qy 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr-----869

Db 2522 ACGGTCCGAACCTCTCTATTATTGTTGATCGGCAAGTGTTTGGAATPACTAAATGGAATCA 2581
Qy 870 -----AspLysGlyPheThrAsn 875
Db 2582 GATAAAATGATTAGAGAGCGATGATTAAAGAGATTGCCGTGATTATGGCAATCAACG 2641
Qy 876 AspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArg 895
Db 2642 CGTATTCGCGCTCTACAGGTGTCGGATTCCAATGGCAATCTCTATTGGGCCATTGGTA 2701
Qy 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
Db 2702 TTCTCTTATGCAACCAATTAATAAATATGAAATATGATGTGCAACAGTTCCAATTT 2761
Qy 914 PheIleGlyThrProPhe 919
Db 2762 AGTATGGAGGTTCTTTC 2779
RESULT 18
US-08-942-046-55
: Sequence 55, Application US/08942046
: Patent No. 6264954
: GENERAL INFORMATION:
: APPLICANT: CHONG, Pele
: APPLICANT: THOMAS, Wayne
: APPLICANT: YANG, Yan Ping
: APPLICANT: LOOSMORE, Sheena
: APPLICANT: SIA, Dwo Yuan Charles
: APPLICANT: KLEIN, Michel
: TITLE OF INVENTION: HAEMOPHILUS OUTER MEMBRANE PROTEIN
: NUMBER OF SEQUENCES: 55
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: 6TH Floor, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/942.046
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/433,522
: FILING DATE: 12-SEP-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: STEWART, Michael I
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-732 MIS:jb
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 55:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2987 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-942-046-55
Alignment Scores:
Pred. No.: 2.08e-08 Length: 2987
Score: 172.00 Matches: 171
Percent Similarity: 35.21% Conservative: 141
Best Local Similarity: 19.30% Mismatches: 344
Query Match: 3.64% Indels: 230

DB: 4 Gaps: 44
US-09-914-168-2 (1-919) x US-08-942-046-55 (1-2987)
QY 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAla 198
DB 377 AGGATACATCGATGAAAAAATCTTCTAATCGCAAGTTTATTATTTCGGTACGACACGACT 436
QY 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp 218
DB 437 GTGTTGGCGCA- - - - -CCTTTGTGGCAAAAGATATTTCGTGCGATCGT 481
QY 219 IleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
DB 482 GTTCAAGGTGACTTAGAACAAACAAATCGACCAAGTTTACCT- - - - - 523
QY 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
DB 524 GTTCGTGCGCGTACGCTGACTGACACATGATG- - -GCTAATATTGTCCGCTCTTTA 580
QY 259 Ile- - - - -GlyGluVal- - - - -AspValIleIleHisAspLeuGlyGluProValTyrIle 275
DB 581 TTCGTAAAGTGGTCGATTCGATGATGTAAGCGCATCAAGAGCGGATGCTGTTGTT 640
QY 276 AspTyrArgAla- - - - -ValGluValArgGlyGluGlyAla- - -Asp 288
DB 641 AGCGTTGGCTTAATCGCATCATTTTCAGATGTTAAATCAAGTAACCTCTGTTATTCCT 700
QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
DB 701 ACTGAAGCACTTAACAAAACCTTAGATGCTAACGGGTTTAAAGTTGGCGATGTTTA- - - 757
QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyr 328
DB 758 - - - - -ATTCGAGAAAAATTAATGAATTTGCCAAAAGTGTAAAAGAGCAGCATATGCA 808
QY 329 PheAspGlyArgTyrLeuAspArgSerValAspValIle- - - - -LeuProAspAsn 345
DB 809 AGTGTAGTGCCTAT- - -AAGCGAACAGTTGAACCTATTGTCAATAGCTACCAAAATAT 865
QY 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
DB 866 CGCCTGAAATTTTAAATCAAAATCAATGAAGATGATGAAGCAAAATTTGGCATCACTA 925
QY 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
DB 926 TTC- - - - -AAGGGAACGAA- - -TCTGTTAGTAGCAGTACATTACAAGAA 967
QY 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr- - - - - 402
DB 968 CAAATGGAATTACAACTGATTCTTGTGGGAAATTTATGGGAAATAATTTGAAGTGCG 1027
QY 403 - - - - -AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
DB 1028 CAATTCGAGAAAGATTGCGAGTCAATTCGT- - - - -GATTAT 1063
QY 419 TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAsp 438
DB 1064 TATTTAATATGCTATGCTATGCAAGCACAATTAATAACGGGATGTTACAGCTAAATGAT 1123
QY 439 GlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp- - - - - 456
DB 1124 GAAAAACAAAAGTTAATGTAACCATTTGATGTAATGAAGGTTTACAGTATGACCTCGT 1183
QY 457 - - - - -GluSerThrLeuGluProValIleGlu 465
DB 1184 AGTCGACGCAATTAGTAATCTGGGAGTATGTCGCCAGGCTTGAACCTTTACTTTCA 1243
QY 466 ThrValGluLeuThrAspGlyIle- - - - -LeuMetAspIleSerProIleGluPheSerAla 484
DB 1244 GCATTACATTAAATGATACTTTCCGCCGTAGTGATATTGCCAGATGTAGAA- - - - - 1294
QY 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504

DB 1295 - - -AATGCAATTAAAGCAAACTTTGGACACCGGTTACGCTACGCCAACCGTAAATTCA 1351
QY 505 MetProAsp- - - - -AspArgValLeuAlaIleAsnHisAspAspGlyValAsn 520
DB 1352 GTACCTGATTTTGTATGATGCAAAATAAAACATTAGCGATAACCTTTGTTGTTGATCGTGA 1411
QY 521 ArgSerIleLeu- - - - -GlyArgIleSerAspAlaValSerAla 533
DB 1412 CGAGCTTTAACTGTTGCGCAACTTCGGCTTTGAAGGAATAACCGTTTCTGCTATGACACT 1471
QY 534 ValAlaArgAlaIleLeuProAspGluSer- - - - -GluAsnGluValIleAspLeu 550
DB 1472 TTACGTCAGGAAATCGCCCAACAAGAGAACTTTGGTATATATTCACAAATTAGTTGAGTTA 1531
QY 551 Pro- - - - -GluArgThr- - - - -AlaLeuAlaAsnArgLysThrPro 562
DB 1532 GGAATAATTCGCTTTAGATCGTACAGGTTTCTCGAAACAGTTCGAAACCGAATTCATCCT 1591
QY 563 AlaAsp- - - - -ValTyrGlnSerLysLysValProLeuTyr 574
DB 1592 ATCAATGCTAGTATGATGAAGTGTGCTGATATATAAGTCAAA- - - - - 1636
QY 575 ValPheValAlaSerAspLysProArgAspGlyGlnIle- - - - -GlyLeuGlyTrpGly 592
DB 1637 - - - - -GAACGTAACACGGTAGTATCAACTTTTGTATTTGCTTACGCT 1678
QY 593 SerAspThrGlyThrArgLeuValThrLysPhe- - - - -GluHisAsnLeuIle- - - - - 608
DB 1679 ACAGAGAGTGTATTACTTATCAAGCAAGTGTAAACAGATAATTTCTTTGGGAACAGG 1738
QY 609 - - - - -AsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
DB 1739 GCGCAGCTAAGTATAGCTGCTACGAAAATGATTATGGTACGAGTGTCAATTTTCGGTTAT 1798
QY 622 SerGlu- - - - -AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
DB 1799 ACCGAGCCCTATTACTAAAGATGCTGTAAGTCTTGGTGAATAATGTTTCTTTGAAAC 1858
QY 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSer 658
DB 1859 TACGATAACTCTAAAGTGATACATCTCTAACTATAAGCGTACGACTTACGGAAGTAAT 1918
QY 659 - - - - -ThrAsnGlyPheAspLeuSerThrArgThr- - - - -LeuGluHis- - - - - 671
DB 1919 GTTACTTTAGGTTTCCCTGTAAATGAAAATAACTCTATTATGTAGGATTAGGTCATACC 1978
QY 672 - - - - -GluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThr- - - - - 686
DB 1979 TATAATAAAATTAGTAAGTTGCTCTAGAA- - - - -TATAACCGTAATTTATATATAT 2029
QY 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGln 706
DB 2030 CAATCAATGAAATTTAAAGTAATGCGATTAAACAAAT- - - - - 2068
QY 707 AspLeuProValAspPheValAsnGlyLysProSerGlnAlaLeuLeuAlaGlyVal 726
DB 2069 - - - - -GACTTTGATTTTCTTTTGGTTGG- - - - - 2092
QY 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
DB 2093 - - - - -AACTATACAGCTTAAATAGAGGCTATTTCCTCCAACT 2128
QY 747 TyrSerLeuGluValGlySerSerGlyLeuVal- - - - -SerAspAlaAsnMet 762
DB 2129 AAAGGGTTAAAGCAAGTCTTGTGCGCAGGTTACTATTCCAGGTTCTCATACAATAATAC 2188
QY 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
DB 2189 TACAACACTAAGTCGACATGTACAGGTTTCTACCCATTA- - - - - 2227
QY 783 AsnArgAlaHis- - - - -GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAsp 800

Db 2228 GACAGAGATCACCTCTGGTGTATCTGCAAAAGCATCTGCAGGATATGCAAAATGGTTTT 2287
Qy 801 AsnPheAsnHisValProTyrArgLeuArgPheAlaGlyGlyAspGlnSerIleArg 820
Db 2288 GGAACACAGCGTTTACCGTCTATCAAACTTATACAGCGGTGGCATCGTTTCATTACGT 2347
Qy 821 GlyTyrAlaHisAspSerLeuSerPro----- 829
Db 2348 GGTTCGTATGTAGTAGTATGGAGCTAACGCAATTTATGCCGAATATGGTAATGGTAGT 2407
Qy 830 -----IleSerAspLysGlyTyrLeuThrGlyGlnValLeu 842
Db 2408 GGTACTGGTACTTTTAAAGAAGTAGTCTGTAT-----GTATGGTGGTAAATGCAATC 2461
Qy 843 AlaValGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp----- 856
Db 2462 GCTACAGCTAGCGCAGAGTAAATTTGCGCAACTCCATTTGTGAGCGATAAGAGCAAAAT 2521
Qy 857 ---LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
Db 2522 ACGGTCCGAACCTCCTATTGTTGATCGGCAAGTGTTCGGAATACTAAATGGAATCA 2581
Qy 870 -----AspLysGlyPheThrAsn 875
Db 2582 GATAAAATGGATTAGAGCGCATGATTAAAAAGATTGCCCTGATATTGGCAAAATCAAGC 2641
Qy 876 AspThrLysIleGlyAlaGlyValGlyValArgTyrPalaSerProValGlyGlnValArg 895
Db 2642 CGTATTCGGCGCTACAGGTCGGATTCCCAATGGCAATCTCCTATTGGGCCATTGGTA 2701
Qy 896 ValAspValAlaThrGlyValLys-----GluGluGlyAsnProIleLysLeuHisPhe 913
Db 2702 TTCCTTATGCCAACCAATTAAAAAATATGAAATGATGATGTCGCAACAGTTCCAATTT 2761
Qy 914 PheIleGlyThrProPhe 919
Db 2762 AGTATTGGAGGTCTTTC 2779

RESULT 19
US-07-689-008-5
Sequence 5, Application US/07689008
Patent No. 5268274
GENERAL INFORMATION:
APPLICANT: Ben-Bassat, Arie
APPLICANT: Calhoun, Roger D
APPLICANT: Fear, Anna L
APPLICANT: Gelfand, David H
APPLICANT: Meade, James H
APPLICANT: Tal, Roni
APPLICANT: Wong, Ring
APPLICANT: Benzimid, Moshe
TITLE OF INVENTION: METHODS AND NUCLEIC ACID SEQUENCES FOR THE
TITLE OF INVENTION: EXPRESSION OF CELLULOSE SYNTHASE OPERON
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: McCutchen, Doyle, Brown & Enersen
STREET: Three Embarcadero Center
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/689,008
FILING DATE: 19910422
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 337,194
FILING DATE: 12-APR-1989

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 496,236
FILING DATE: 23-MAR-1990
ATTORNEY/AGENT INFORMATION:
NAME: Murphy, Lisabeth Feix
REGISTRATION NUMBER: 31547
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 393-2000
TELEFAX: (415) 393-2286
TELEX: 340817 MACPAG SFO
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3957 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-689-008-5

Alignment Scores:
Pred. No.: 8.13e-07 Length: 3957
Score: 159.50 Matches: 201
Percent Similarity: 31.65% Conservative: 131
Best Local Similarity: 19.16% Mismatches: 360
Query Match: 3.37% Indels: 357
DB: 1 Gaps: 51

US-09-914-168-2 (1-919) x US-07-689-008-5 (1-3957)

Qy 43 ProAlaHisAspThrAlaIle----- 49
Db 490 CCGCGCCATCTCCGCGGTGGAATACTACCAGACCATGCGGGCGTGCAGCCAGTGG 549
Qy 50 AsnGlnAlaLysAlaGlyAsnProValLeuLeuThr---ProGluGlnIleGlnAla 68
Db 550 GACCAAGCGCGCGCGGGTTCGCGGTGCTCAACCCGAGGATTACCGCGCC 609
Qy 69 ArgLeu-----AsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeu--- 84
Db 610 CAGCTCGCCTTTGCCAGCGCCCTGACCTATAATACCTCAGCCGATGGAAGCCTGACC 669
Qy 85 -----AspValValAsnPheAspGlnSerProIle---SerArgIleGlyGlu 100
Db 670 CGGCTCAAGGATCTCCAGTCTTCGCGAGCCAGCGCGTCCGAGCGCCGCCCGCGCG 729
Qy 101 GlnSerProLeuGlyLeuAsp---MetSerValIleGluGluThrThrProLeuSer 119
Db 730 CAGTCTACCGCCAGACCCCTGAGCTGGCTGCGGTCAATCTCGAGACGAGCCCTCATG 789
Qy 120 LeuGluGluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIle 139
Db 790 GAGCAGTGGCTTTCCGCC-----CACCCCAATGAT----- 819
Qy 140 ProGluTyrGlnGlyGluGlnProAsnSerGluValValValProProThrLeuGluPro 159
Db 820 -----ACCGCGCTGCGAGCATATGCTCCACCCCGCGTGTGTCGCGCG 864
Qy 160 GluLysProGlyLeuIleLysArgLeuTyrAlaAtgLeuPheAsnAspGlyValAsnLys 179
Db 865 GACAAGCGCGGTTCGCGCGCAG----- 888
Qy 180 ValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAlaIle 199
Db 889 -----GCAGTTTACCAGACGCTTACCGCGCGTCTTTCGCGCGAGCC 930
Qy 200 GlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIle 219
Db 931 GAGCAGTCTTTCAG----- 945
Qy 220 ThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeu 239
Db 946 -----TCGGCGTTGCAGATCAATTC----- 966

QY 240 ValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIle 259
Db 967 -----CATGATCTGAT-----TCGCTT 984
QY 260 GlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAla 279
Db 985 GGTGGCATGGGCTCGTAGCATGGCGAGGGCGATACCGCG-----GAGCGCGCGCGC 1038
QY 280 ValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAlaGluVal 299
Db 1039 TATTTTCAAGAGCGATGGCGCGCGACCCCAAG-----ACCGCGCATCGCTGG 1086
QY 300 ProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeuIle 319
Db 1087 CGCGCGCGCTTCGGGGCATGGCGCTACGCGCGAGTATGCTTCGGTTCCGTCAGTTGATT 1146
QY 320 GluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrp----- 333
Db 1147 -----GCGGCCCATCAATATACCGAGGCGCAAGCAGCAGCTTGCACGCTGGCC 1194
QY 334 -----LeuAspArgSerVal 338
Db 1195 CGCGACCGCGCAGTATACTGGCGGACCTCATGCTGGCGGACCTGCAGCGCTCGACC 1254
QY 339 AspValIle----- 341
Db 1255 GCCCAGATTGGCGCGCGCAGGAGGAATATCGTGCATCCTGTCGCTGAGGCCCAATAAC 1314
QY 342 -----LeuProAspAsnThrAlaAspVal 349
Db 1315 CAGTTGGCCCTCATGGGCTGGCCGGGTAGACATGGCGCAGGGCAACACGGCGGAAGCA 1374
QY 350 ---SerLeuIleTyrAspThrGlyThrGlnTyr-----ArgPheAspGluValValPhe 366
Db 1375 CGCCAGCTCCTGCTCGGCTGTCGGCCGCAATATGCAAGCCAGGTGGCGAGATGAGGTT 1434
QY 367 PheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLys 386
Db 1435 TCGGCGCTGATGGCGGCTGCTCCAGACATCGGATTCAGCGCGCAAGTTTCCATCTG 1494
QY 387 ArgGluLeuLeuGluGlnLeu-----LeuThrValAsnMetGlyGluAla 401
Db 1495 CGCGAAGCGATGGCCAGGCCCCAGCCCGTGCATGCGCGCATCAAGCTTGCCAATGCG 1554
QY 402 TyrAsnLeuGln-----AlaValArgAlaLeu----- 410
Db 1555 CTGCAGCAGCAGGGGAGGTGGCCGAAGCGCGGCGGTGATGACGCCATCCTGGCGAAT 1614
QY 411 -----SerAsnAsp 413
Db 1615 CCCGTACCGCGCAGGACCGCGCGGTATCCTTTTATACCTATGTTAGTGGCAATGAT 1674
QY 414 LeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGlu 433
Db 1675 CGCATGACCCCGCAGCTCTGGCTGTCTGCTCGGCTCGGATTAATCTCTGTCATCCGT 1734
QY 434 GlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArgThrGluProAla 453
Db 1735 TCCATCGCC---GAGGAATGGAATCAAGCAGGATCTGGCCAGCGCCCTGTCCATGGT 1791
QY 454 GlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIle 473
Db 1792 TCCAACCGGTGGCTGATCCCGGAGGCGCTGACCCAGCGCTGATCCGACCGCGCGCGC 1851
QY 474 LeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsn 493
Db 1852 GGCCTGGCTGGTGCACCTTCCCGCAGCGTGGCGACATGGTGCATGCCCGCCATGGCA 1911
QY 494 LeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspArgVal----- 510
Db 1912 CTGCGTATCCGCTCGAGCGCGCAGCATGATCTCTCGCGCCGACCGCGCTGCTCTATGCC 1971
QY 511 -----LeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIle 527

Db 1972 ACCGAATACATCAAGATCAGCAACCGGTCGGCTGGCGGG---CTGTCGCCCGCGTGG 2028
QY 528 SerAspAlaVal---SerAlaValAlaArgAlaIleLeuProAspGluSerGlu----- 544
Db 2029 GCGGATGGCAGCGGCTCGGCTACAGGAAGCGGTTGCTGCCGAGCAGCTGCACAGCGCTC 2088
QY 545 -----AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThr 561
Db 2089 CAGCAACTGGCGCATGGGATCTCGGTGGCGAGTCTGCTCAACACCGCTGGCGAC 2148
QY 562 ProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLys 581
Db 2149 CAGGCGCAGCGCTATGATCATCTGCCCGCGCTGCAGCGCGACCCCGAGCGGCACATCG 2208
QY 582 ProArg-----AspGlyGlnIleGlyLeuGlyTyrGly 592
Db 2209 CCCAAGCTGGGCTGGCGCGCTGTATATGGCCAGCGCAAGCGCGCAAGCGCTCGAG 2268
QY 593 SerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGly 612
Db 2269 ATCGAC-----CTTGGCGTGTGGCGCCACAACCCCGCAGGACCTTTCATGCG 2313
QY 613 TyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyVal-----LysLeu 630
Db 2314 CGACAGGCTGGCTGCGAGCGCGGTCAACAGCAGCACCAACAGCCTTGCCACACCGCTT 2373
QY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
Db 2374 GCCATGATGGCTGCAGGAAGCCGATGATGCCCTGGCTGGCTGGCGCATGGCGGTG 2433
QY 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670
Db 2434 GCTCAGCAGCGCATGGCCAC-----GGCAGCGCAGCATCGAG 2472
QY 671 HisGluIleSerArgSerIleIleGlnAsnGlyTyrTrpAsnArgThrTyrSerLeuArg 690
Db 2473 -----GATCTGGCGCGCGCTATGATGCTGCGC 2499
QY 691 TyrArgLeu-----AspLys- 695
Db 2500 CTGCAGCAGGTTCGAGGCGCGCGGCGCTGTCGGCGCGGTGCTGGCGAGGAGATGCG 2559
QY 696 LeuLysThrGlnAlaProProGluThr-----Trp 705
Db 2560 CTT-----GCTCGCGCTCGACCAACCGCTTCGCCCGCGCTGGCTACGCCACCGAG 2610
QY 706 GlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAla----- 721
Db 2611 ACGGAATTCGGCGCTGTGACGGGTGCTCTACAGCGCGCGCGGCGCATCGCCCGAT 2670
QY 722 -----LeuLeuAlaGlyValAlaValHis---LysThrValAlaAspAsnLeu 736
Db 2671 ACCTCGGACCATGCTCTCTCTCCATCGCAGCGCAGATCCGACAGCTGCTGAGACCTT 2730
QY 737 -----ValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySer 753
Db 2731 GCCCTTTCATCATGTCGTCGCGGTTCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 2784
QY 754 SerGlyLeuValSerAspAlaAsnMetAlaIleAla-----ArgAlaGly 768
Db 2785 ATGGCGCGCTGACGGAAGCAACATTCCTCATGTCGGCGCGCTGCGCGCTGCGCGCGT 2844
QY 769 IleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMet 788
Db 2845 GCTTCGCGC-----CTG 2856
QY 789 ThrGlyGlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsn----- 803
Db 2857 ACCTTCTCATCAGCCACCATGATCTCGTCGGCAACCTCAACACCGGTTCCCTCTAT 2916
QY 804 HisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyr--- 822

Db	2917	GATGTCCG	-----CGTTAT-----GGCAGATGATGGCGTCGAGGCATATAAC	2961
Qy	823	AlaHisAspSerLeuSerProIleSerAspIysGlyTyrLeuThrGlyGly	-----839	
Db	2962	CAGTACGATAGC	-----TATACCAACGCGGCGCAGGACCAG	2997
Qy	840	GlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeu	859	
Db	2998	CAGCGCATCGCGCTCGCACGCGCAG	-----3024	
Qy	860	AlaValPheGlyAspIleGlyAsnAlaTyrAspIysGlyPheThrAsnAspThrLysIle	879	
Db	3025	-----	-----GCCGGTTGGCGCGGATGTCAGTTT	3051
Qy	880	Gly	-----AlaGlyValGlyValArgTTrpAlaSerProValGlyGlnVal	894
Db	3052	GGCAATAGCTGGTGGCGCGCATGTGGT	-----GCGTCGCCCATCGGCTTCCCC	3102
Qy	895	ArgValAspValAlaThrGlyValIys	903	
Db	3103	ATCACCAAGCTGTGGCGGCTCGAG	3129	
RESULT 20				
US-07-689-008-1				
: Sequence 1, Application US/07689008				
: Patent No. 5268274				
: GENERAL INFORMATION:				
: APPLICANT: Ben-Bassat, Arie				
: APPLICANT: Calhoun, Roger D				
: APPLICANT: Fear, Anna L				
: APPLICANT: Gelfand, David H				
: APPLICANT: Meade, James H				
: APPLICANT: Tal, Rony				
: APPLICANT: Wong, Hing				
: APPLICANT: BenZiman, Moshe				
: TITLE OF INVENTION: METHODS AND NUCLEIC ACID SEQUENCES FOR THE				
: NUMBER OF INVENTION: EXPRESSION OF CELLULOSE SYNTHASE OPERON				
: NUMBER OF SEQUENCES: 15				
: CORRESPONDENCE ADDRESS:				
: ADDRESSEE: McCutchen, Doyle, Brown & Enersen				
: STREET: Three Embarcadero Center				
: CITY: San Francisco				
: STATE: California				
: COUNTRY: USA				
: ZIP: 94111				
: COMPUTER READABLE FORM:				
: MEDIUM TYPE: Floppy disk				
: COMPUTER: IBM PC compatible				
: OPERATING SYSTEM: PC-DOS/MS-DOS				
: SOFTWARE: PatentIn Release #1.24				
: CURRENT APPLICATION DATA:				
: APPLICATION NUMBER: US/07/689,008				
: FILING DATE: 19910422				
: CLASSIFICATION: 435				
: PRIOR APPLICATION DATA:				
: APPLICATION NUMBER: US 337,194				
: FILING DATE: 12-APR-1989				
: PRIOR APPLICATION DATA:				
: APPLICATION NUMBER: US 496,236				
: FILING DATE: 23-MAR-1990				
: ATTORNEY/AGENT INFORMATION:				
: NAME: Murphy, Lisabeth Feix				
: REGISTRATION NUMBER: 31547				
: TELECOMMUNICATION INFORMATION:				
: TELEPHONE: (415) 393-2000				
: TELEFAX: (415) 393-2286				
: TELEX: 340817 MACPAG SFO				
: INFORMATION FOR SEQ ID NO: 1:				
: SEQUENCE CHARACTERISTICS:				
: LENGTH: 9540 base pairs				
: TYPE: NUCLEIC ACID				
: STRANDEDNESS: single				
: TOPOLOGY: linear				

QY 555 AlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeuTyr 574
Db 3999 -----GTGAATCCAGAACTCAACAACTGAAACCTTTACATCTCGTGCC----- 4040
QY 575 ValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAsp 594
Db 4041 TTTCTTTGTAGACAGCATGATCCCGT-----GGTGTCTCTTTCTTTTCCCGT 4088
QY 595 ThrGlyThrArgLeuValThrLysPheGluHis---AsnLeuIleAsnArgAspGlyTyr 613
Db 4089 ACAGGTAACGAGTGAAGTGAAGAAAGAACTCATGCAACATCGTC----- 4133
QY 614 GlnAlaGlyAlaGluLeuArgLeuSerGluAspLysGlyValLysLeuTyrAlaThr 633
Db 4134 -----TTT 4136
QY 634 LysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGlu 653
Db 4137 AAACAAATGGATTCATCTTTGGAGAACCACTTGCTCCAAATATTT-----TTGACC 4187
QY 654 ValPheGlyHisSerThrAsnGlyPheAspLysSerThrArgThrLeuGluHisGluIle 673
Db 4188 ATCTATATTCAACACAGAGAGCTTCTCTTAGCTTAATGGGAGCAAGTAGGAGAA 4247
QY 674 SerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThrTyrSerLeuArgTyrArgLeu 693
Db 4248 GAATTTAACTTGGCTCTTAAC-----TCACTTGATTACCGGTACA 4286
QY 694 AspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheVal 713
Db 4287 GATGCGAGTCAACTGGTGCTTCTCCAGAACCATACGAAATTTGATTATGATGCTTA 4346
QY 714 Asn 714
Db 4347 AAT 4349
RESULT 22
US-09-453-702B-50
Sequence 50, Application US/09453702B
Patent No. 6365723
GENERAL INFORMATION:
APPLICANT: Blattner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 50:
SEQUENCE CHARACTERISTICS:
LENGTH: 38584
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-09-453-702B-50
Alignment Scores:
Pred. No.: 0.015 Length: 38584
Score: 136.50 Matches: 183
Percent Similarity: 32.36% Conservative: 126
Best Local Similarity: 19.16% Mismatches: 360
Query Match: 2.89% Indels: 286
DB: 4 Gaps: 42
US-09-914-168-2 (1-919) x US-09-453-702B-50 (1-38584)
QY 53 LysAlaGlyAsnProProVal-LeuLeuThrProGluGlnIleGlnAlaArgLeuAsnAl 72
Db 22727 AGACAGCGCAACCGCCAGAACTCTTCTTCTGAGTTGCTTCTGAGCTGAGCGCA 22786
QY 72 aAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspGln 92
Db 22787 GCTGGGGGTA-----AGCAGTGTCTGCAAGAGTGTAT-----AT 22819
QY 92 nSerProIleSerArgIleGlyGlnSerProProLeuGlyLeuAspMetSerVal 112
Db 22820 CCGCAAAATGAGCAGCGCAAGAAAGCCAGCATTCCTCGGTCTGAAAGATGCGAGCG 22879
QY 112 eGluGluThrProLeuSerLeuGluGluPheAlaGlnGlu----- 127
Db 22880 CCGTCAGCAGATAGCGGTTCTGATAGTGAACCTGGCCCGCGGCAATTTAGTGGCTGAG 22939
QY 127 ----- 127
Db 22940 GGGATCCGGAATAAGCGTGGCTAACCGTCCGGGTGATAGACACACTGATGTACCGAA 22999
QY 128 -----SerThrGluMetGlyIleAsnPr 135
Db 23000 AGCATGATCCAGAGCGGTTATTGGCGGTATTGGCGGCTCTCTATGCTCTGGGTAA 23059
QY 135 oAsnAspTyrIleProGluTyrGlnGlyGluGlnProAsnSerGluValValValProPr 155
Db 23060 AGCTGTATGACGGTCAGAGAGGAGGGGAGAAATTTAACCGCAGCTGCTGCTGACGGG 23119
QY 155 oThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAs 175
Db 23120 GCATTATGCCGGAGTCACTGCCGGGAGCTGTGAGCAGTCACTGCTGCTATTTCCGGGAA 23179
QY 175 pGlyValAsnLysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyG 195
Db 23180 TGGTATCACGCAA-----CATGCTGACGCGCGTGGCT 23212
QY 195 uThrSerAlaIleGlySerHisGlnLysThrGluPro-----TyrAlaAsnIl 212
Db 23213 GGCTCAGTGTGGGAGTGGTGCATTTCTGTGAAACCATATCGGTATGTTGGCGAGAGC 23272
QY 212 eLysAlaAlaLeuGluAspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIlePr 232
Db 23273 TGGCGCACAGATGGAGCGCATCGGTTGGCCAGTCTGTCAGCGATACCAATAGTTTAA 23332
QY 232 oArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLe 252
Db 23333 CGCGCTGAAGGATGATCTGTAAATGCCGCAAGGCT----- 23369
QY 252 uSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGluPr 272
Db 23370 -----CTGCAATGAGCTGCTATTTCTTACTGCCAC 23401

Qy 272 oValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAspLys----- 290
Db 23402 TCAGCTTTGAGCAGATACCGCTCTTGGGAGCAGGGCGGTCCAGTGATGTCGACGGAT 23461
Qy 291 -AlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisG1 310
Db 23462 AGCCATCTGCACTGGCAGGAGAAACCGGTGCGGTACTGCGGATATT----- 23510
Qy 310 yLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAs 330
Db 23511 -----GATAATAACCTCAATGCCCTTGGCAGTACGCTGAAGTATCTGTC 23554
Qy 330 pGlyArgTyr-----LeuAs 335
Db 23555 TGATTTCTGGAGTCGTTCTGGGATCGGCCATGAATATTGGTCTGAAAGACTCGCTGGA 23614
Qy 335 pArgSerValAspValIle-----LeuProAspAsnTh 346
Db 23615 TGAACAGATTGCCCTTTACAGAGAGAAGTGTGCGGGCGGAAAAAGACTCCCTGGACGGC 23674
Qy 346 rAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPh 366
Db 23675 ATCATCTCTCAGGTTGAATACGAT-----CAGCAGGCTCTTACGAT----- 23717
Qy 366 ePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLy 386
Db 23718 -----CTTCAGGAGAAAAAACGCCAGAGGATTTGCGGATGCAAAAGACGAGCAGA 23770
Qy 386 sArgGluLeuLeuGluGln-----LeuLeuThrVa 396
Db 23771 CGGGAATATACAGGAGCAACAGAACGCCGTAATGCTGAAATGCTGCATCGACCCGGAT 23830
Qy 396 lAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAl 416
Db 23831 GAATGAAACGGAACAGCAGCAGCATCAGCGTGAATACCGCTATTAT-----GC 23881
Qy 416 aThrArgTyrPheAsnMetValAsnThrGlnIleValIlePheProGluArgGluGlnIleG1 436
Db 23882 CATGCAGTACGCGCATCAGGCTGTTCAGGATGCGCG-----ATACA 23923
Qy 436 nAsnAspGlnValSerPheGluGln-----SerSerSerArgThrGluProAl 453
Db 23924 ACGTGAATATGACGTTACGAGAAAGCCCTGGCATCCCGTTAAGAAAAAACACCCGGAAC 23983
Qy 453 aGlnValAspGluSerThr-----LeuG1 461
Db 23984 CCGTAATGATGAGCCACCCCGTTATTGCTCAGTACAGTACAGCAACAGGCACAGGTGGA 24043
Qy 461 uProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleG1 481
Db 24044 AGGACAGATTGCTGCTGCCACAGACAGTACAGGCAATT----- 24080
Qy 481 uPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgH1 501
Db 24081 -----GCCACGGAAGGATGACAGAGCGCGTAAACAGCTTCGCTGTCGACGACGCG-- 24134
Qy 501 sLeuTyrAspMetProAspArgValLeuAlaIleAsnHisAspAspGlyValAsnAr 521
Db 24135 -ATCAGCGACCTGGACGGGAAAAAAGTACGCGCA-----CATGAA-----AA 24175
Qy 521 gSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAs 541
Db 24176 GAGTGTCTGGCGCGTAAAGATGAATCACTGATTCAGGCACTGACG-----CTGCTGGA 24226
Qy 541 pGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysTh 561
Db 24227 TGTAACACAGCAGGAG-----CTTCAGAAACACAGCGCACTCAACGAGCTGAAG-- 24275
Qy 561 rProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLy 581
Db 24276 -----AAAAAACAAATTCAGCTGACCCAGTCACTGCTGAAGAGA 24316

Qy 581 sProArgAspGlyGln-----IleGlyLeuGlyTyrPglySe 593
Db 24317 CGCGCTCAGCGTCAGCAACATGACCTGATATGCCACGGTGGGTATG----- 24365
Qy 593 rAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTy 613
Db 24366 -----GGTGATCAGCAGCAGCAGCATATCAGTACAACCTGAGTCTTCGCGAATAA 24418
Qy 613 rGlnAlaGlyAlaGlu-----LeuArgLeuSerGluAspLysLysGlyValLysLeuTy-- 631
Db 24419 CCAGCAACAGCTGAGCAGCTTGAGCGGGATAGTAGCAGAGAAAGGACA-----TATAA 24472
Qy 632 -----AlaThrLysProLeuSerHisProLeuAsnAspGlnLeuAr 645
Db 24473 CACGGATGACTACAGAAAGCGCAGCAGCGCTGACGAGAGCTGAAACCGACAACATAA 24532
Qy 645 gAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSe 665
Db 24533 TGAGAAATCGCGTTACTGTGCAACAGCTT----- 24560
Qy 665 rThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnAr 685
Db 24561 -----GAAGTTCTGACAG-----GGTAACCTGG--AA 24583
Qy 685 gThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTr 705
Db 24584 AAACGGAGTCTCGCT-----GCATT 24604
Qy 705 pGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaG1 725
Db 24605 TCAGGATTTTACCCTGGGATGAGATATACGGCAGAAACAGCAGAA----- 24650
Qy 725 yValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgG1 745
Db 24651 ---CAGGTGTCTCGTCAGCCTTCAGCAACATGGGAATGCCCTGGCAACTTTTGTCTAC 24706
Qy 745 nArgTyrSerLeuGluValGlySer-----SerGlyLeuValSerAspAlaAsnMetAl 763
Db 24707 TACCGGCAACTCAATTTCAAACTCTTCACTCTCTCTGTGTGTGATATGCGGAAAAAT 24766
Qy 763 aIleAlaArgAla-----GlyIleSerGlyValTyrSerPheG1 776
Db 24767 CCTGGCGCAGGACCACTGATGAATCGAATAAAGGGATTCGACGTACTGGGATTGGA 24826
Qy 776 yAspAlaAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTy 796
Db 24827 TCTCAGCAGCCTTTCCCTGTAAT-----GCCAATGGGGGATTTATCAGTCTGC 24874
Qy 796 rIleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAla----- 813
Db 24875 TGATTTGAGTCGTATACAGTGCACGGTGTAAACGTCGACGCTTTTGTCTTTTGCAAA 24934
Qy 814 -----GlyGlyAspGlnSerIle-----ArgGlyTyAl 823
Db 24935 AGCGCGGGTGTATGGGGGAGCGGACCTGAAGCCATTCTGCCATTCGCTGCTGGTGC 24994
Qy 823 aHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAl 843
Db 24995 TGACGGTAAGCTGGGGTGTGCGGAT-----ATTGGGGGTTCAGGTATGCG 25042
Qy 843 aValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheG1 863
Db 25043 GATGTTTTTCCCGCGGATCAACATCGAGATCAATAACGAT-----GGCAGCAACGG 25093
Qy 863 yAspIleGlyAsnAla-----TyrAspLysGly 872
Db 25094 GCAGATAGTCCGCTGCCCTGAAAGCGGTTTATGACCTCGGG 25136

RESULT 23

US-08-621-944A-2

; Sequence 2, Application US/08621944A

; Patent No. 6440425

; GENERAL INFORMATION:

APPLICANT: SASAKI, Ken
 APPLICANT: HARKNESS, Robin E.
 APPLICANT: LOOSMORE, Sheena M.
 APPLICANT: KLEIN, Michel H.
 TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER
 MEMBRANE PROTEIN OF MORAXELLA
 NUMBER OF SEQUENCES: 10
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sim & McBurney
 STREET: Suite 701, 330 University Avenue
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5G 1R7
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/621,944A
 FILING DATE: 26-MAR-1996
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/478,370
 FILING DATE: 07-JUN-1995
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I.
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1038-587
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5976 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-621-944A-2

Alignment Scores:
 Pred. No.: 0.000655 Length: 5976
 Score: 136.00 Matches: 191
 Percent Similarity: 32.95% Conservative: 128
 Best Local Similarity: 19.73% Mismatches: 351
 Query Match: 2.88% Indels: 298
 DB: 4 Gaps: 49

US-09-914-168-2 (1-919) x US-08-621-944A-2 (1-5976)
 Qy 22 ProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnProAlaAsnIleIleAsnHis 41
 Db 1357 CCATATTGGATAAAACAACTTAAAGTGGGTAGTGTTCATATACCATAGACAATGGC 1416
 Qy 42 ValProAlaHisaspThrAlaIle---AsnGlnAlaLysAlaGlyAsnProValLeu 60
 Db 1417 ATTGATCGAGTATAAAAGATCAGTAATCTTGCCAAAGGTAGCAGTGTACGATGCG 1476
 Qy 61 LeuThrProGlnIleGlnAla-----ArgLeuAsnAla---AlaGlyLeuAsn 76
 Db 1477 GTTACCATCCACAGCTCAAGCGCCAGCCCTACTTTAAACGCGAGCGCTGGCATCAGT 1536
 Qy 77 AlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspGlnSerProLysSer 96
 Db 1537 GTCACACCTACTGAAATATACAGTTGATGCTAAGAGTGGCAATGTTACCCGCCCACTTAC 1596
 Qy 97 ArgIleGlyGlnSerProProLeuGlyLeuAsp-----MetSerVal 111
 Db 1597 AACATTGGCTGAAACACCCAGCTTAACAGTGGCACTAGTGATGATAAAATTAGTGT 1656
 Qy 112 IleGluGluThrThrProLeuSerLeu-----GluGluLeuPheAlaGlnGlnSerThr 129

Db 1657 AAGGGTAGTGGTACGAACAATAAGCTTAGTTACCGCGCAACATTTTGGCAGCTATCTAAT 1716
 Qy 130 GluMetGlyIleAsnProAsnAsp-----TyrIleProGluTyrGlnGly 144
 Db 1717 GAAGTCATCGAACCGGTGACAGAGTGTCTACAAAGCTTTTACCGGTTAAAGAAGACAGCAT 1776
 Qy 145 GluGlnProAsnSerGluValValProProThrLeuGluProGluLysProGlyLeu 164
 Db 1777 GATGACGCCAACGCTATCACCGTGGCTAAAGATACGACAAA-----1818
 Qy 165 IleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLys 184
 Db 1819 -----AATCGCGCGCAGTCAGCATCTTAAACTCAAA 1851
 Qy 185 AlaLys-----PheTyrGlnSer 190
 Db 1852 GGTAACAAACGGCTTAACGGTGTCTACCAAAAAGATGGTAGGTTACTTTGGGCTTTAGC 1911
 Qy 191 SerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAla 210
 Db 1912 CAAGATACGGTCTGACC-----ATTGGCAAAAGACCCCTAAACACGATGCG---TTG 1962
 Qy 211 AsnIleLysAlaAlaLeuGluAspIleThrGlnGlnSerAla-----Met 225
 Db 1963 ACTGTTAAAGATACCAACGACAAATC---CAAGTCGGTCTAATGCGCATTTAAATTTACT 2019
 Qy 226 AspLeuAsnGlySerIlePro-----ArgLeuArgGlnThrAlaLeuValAlaAlaArg 243
 Db 2020 AATGTGAATGGTAGTAATCAGGTACTGGCATTTGCAAAATACCGCTCGCATTTACACAGAT 2079
 Qy 244 AlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAsp 263
 Db 2080 AAAATTGGCTTTGCTGTTCTGATGGTGCAGTT-----2112
 Qy 264 ValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluVal---282
 Db 2113 -----CATACAAACAAACCT---TATCTTGATCAAGACAAAGCTAAGATTGGC 2157
 Qy 283 -----ArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAla 296
 Db 2158 AATGTTAAGATTACCAACACATGGCATTAACGCGAGTGGTAAAGCCATCACAGGCTGTCC 2217
 Qy 297 AspGluValProLeuLeu-----302
 Db 2218 CCAACACATGCTTAGCATTTGCCGATCAAAAGTACGCCCAACATAGAACTGGCCAAATCAATC 2277
 Qy 303 -----IleGlyAspValPheHisGly---Lys 311
 Db 2278 CAAGACAAGACAATCCAACGCTGCCAGCATTAATGATATATATAATACAGCTTTTAAAC 2337
 Qy 312 TyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGly 331
 Db 2338 CTAAAAATAATAACACCCCATTCGCTTTGCTCC-----2373
 Qy 332 ArgTyrLeuAspArgSerValAspValIle-----LeuProAspAsnThrAlaAspVal 349
 Db 2374 -----ACTATGACATTTGTTGCTTTGCAATGGCAATGCCACCCACCACC 2418
 Qy 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhe---PheThr 368
 Db 2419 ACAGTAACCCATGATACCGCTAAC-----AAAACCACTAAAGTGGTATATGATGCTGAAT 2472
 Qy 369 IleAspProLysThrAsnGlnLeu-----ThrThrAspProAspLysLeuProValLys 386
 Db 2473 GTGGATGATACACCATTCATCTAACAGGCACTGATGACATAAAAACTTGGCGCTCAAA 2532
 Qy 387 ArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuAla 406
 Db 2533 ACCACCAACGTGAAC-----2547
 Qy 407 ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyr---PheAsnMetValAsnThr 425

Db 2548 ---AAAACAAGTCTAATAGTGAATAACAGCAACTAATTTAATGTTAACTCTAGTGATGAA 2604
Qy 426 GluIleValPheProGluArgGluGlnIleGlnAsn---AspGlnValSerPheGluGln 444
Db 2605 GATGCCCTTGTAAACCCCAAGACATCGCCGAAATCTAAACACCTAGCCCAAGAAATT 2664
Qy 445 SerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIle 464
Db 2665 CACACCACCAAGGACACACACACCGCCCTACAAACCTTTACCGTTAAAAAGGTAGAT 2724
Qy 465 GluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAla 484
Db 2725 GAAATAATAATAGCTGATGACGCCCAACGATCAACCGTGGT-----CANAAGAACCCA 2778
Qy 485 SerAsnLeuIleGlnAspLysLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
Db 2779 AATAAT-----CAAGTCACACCTTAACACTCAAAGGTCAAAACGGTCTTAAT 2826
Qy 505 MetProAspArg-----ValLeuAlaIleAsnHisAspAspGlyValAsn 520
Db 2827 ATTAACCGCACAATAATGGTACGGTTACCTTTGGCATTAACACCAACAGCGGTCTTAA 2886
Qy 521 ArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuPro 540
Db 2887 -----GCCGGCAAAAGCACCCCTAAAC 2907
Qy 541 AspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLys 560
Db 2908 GACGGTGGC-----TTGCTATTAATAAACCCCACTGGTAGCGAACAATCCAA 2955
Qy 561 ThrProAlaAspValTyrGlnSerLysValProLeuTyrValPheValAlaSerAsp 580
Db 2956 GTCGGTGTGATGGGTGAAGTTGCCAAGTT----- 2988
Qy 581 LysProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuVal 600
Db 2989 ---ATAATAATGGTGTGTAGTGCTGGC-----ATTGATGGCACAACCTCGCAT 3036
Qy 601 ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArg 620
Db 3037 ACCAGATGAATGGCTTTACTGGGACTAATGGC----- 3072
Qy 621 LeuSerGluAspLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeu 640
Db 3073 ---TCACCTTGATAAA-----AGCAAAACC----- 3093
Qy 641 AsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHis---SerThr 659
Db 3094 -----CACCTAAGCAA 3105
Qy 660 AsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGln 679
Db 3106 GACGGCATTAACGCAGGTGTTAAAGATTACCAACATTCAATCAGTGCAGATTGCCCAA 3165
Qy 680 Asn-----GlyGlyTyrAsnArgThrTyrSerLeuArgTyrArgLeu 693
Db 3166 AACACCCATGATGCTGTGACAGCGGC-----AAGATTATGATTTAAAAACCCGAAC 3219
Qy 694 AspLysLeuLysThrGlnAlaProGluThrTyrGlnAspLeuProValAspPheVal 713
Db 3220 GAA----- 3222
Qy 714 AsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys---ThrVal 732
Db 3223 ---AACAAATCAGCAGTACTGCCAAAACAGCACAAACCTCATACAGAAATCTCAGTA 3279
Qy 733 AlaAsp-----AsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
Db 3280 GCAGATGAACAGGTAATAACTTTACGGTTAGTAACCCCTTACTCCAGTTATGACACCTCA 3339
Qy 747 TyrSerLeuGluVal-----GlySerSerGlyLeuValSerAspAlaAsnMet 762
Db 3340 AAGACCTCTGTCTATCATCCTTTGTCAGGTGAAACGGGATTACCAACCAAGGTAATAAA 3399

Qy 763 AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer 782
Db 3400 GGTGTGGTGGTGGGCATT-----GACCAACCAAGGCCTTA 3438
Qy 783 AsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrSerAspAsnPhe 802
Db 3439 ACCAGCCCTAAGCTGACCGTGGTAAATAATAATGCAAGGCAATGTCATTGACAGCCAA 3498
Qy 803 AsnHisValProTyrArgLeuArgPheAlaGlyGlyAspGlnSerIleArgGlyTyr 822
Db 3499 AAT-----GGTCAAAATACCATCACAGGACTA 3525
Qy 823 AlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlnValLeu 842
Db 3526 ACACACACTCTAGCTAATGTTACCAATGATAAGGT----- 3561
Qy 843 AlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAsp-----LeuArgLeu 859
Db 3562 AGCGTACGCACACACAGACAGGCG---AATATAATCAAGACGACGACAAACCCGTGCC 3618
Qy 860 AlaValPheGlyAspIle-----GlyAsnAlaTyr 869
Db 3619 GCACGATTTGTGTGTGCTAAGCCGAGCGCTTTAACTTCAAGGCAATGCTGAAGCGTT 3678
Qy 870 AspLysGlyPheThrAsnAspThr 877
Db 3679 GACTTGTCTCCACTTATGACACC 3702
RESULT 24
US-08-945-567D-2
; Sequence 2, Application US/08945567D
; Patent No. 6448386
; GENERAL INFORMATION:
; APPLICANT: SASAKI, Ken
; APPLICANT: HARKNESS, Robin E.
; APPLICANT: LOOSMORE, Sheena M.
; APPLICANT: CHONG, Pele
; APPLICANT: KLEIN, Michel H.
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF
; TITLE OF INVENTION: MORAXELLA
; FILE REFERENCE: 1038-745 MIS
; CURRENT APPLICATION NUMBER: US/08/945,567D
; CURRENT FILING DATE: 1996-04-29
; PRIOR APPLICATION NUMBER: 08/431,718
; PRIOR FILING DATE: 1995-05-01
; PRIOR APPLICATION NUMBER: 08/478,370
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: 08/621,944
; PRIOR FILING DATE: 1996-03-26
; PRIOR APPLICATION NUMBER: PCT/CA96/00264
; PRIOR FILING DATE: 1996-04-29
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 5976
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
US-08-945-567D-2

Alignment Scores:
Pred. No.: 0.000655 Length: 5976
Score: 136.00 Matches: 191
Percent Similarity: 32.95% Conservative: 128
Best Local Similarity: 19.73% Mismatches: 351
Query Match: 2.88% Indels: 298
DB: 4 Gaps: 49

US-09-914-168-2 (1-919) x US-08-945-567D-2 (1-5976)

Qy 22 ProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnProAlaAsnIleIleAsnHis 41
Db 1357 CCATATTGTGATAAAAAACAACCTTAAAGTGGGTAGTGTTCGAATTACCATAGCAATGCG 1416

QY 42 ValProAlaHisAspThrAlaIle-----AsnGlnAlaLysAlaGlyAsnProProValLeu 60
Db 1417 ATTGATGCGAGTAATAAAGATACGTAATCTTGGCAAGAGTAGCAGTGTACACGATCG 1476
QY 61 LeuThrProGluGlnIleGlnAla-----ArgLeuAsnAla---AlaGlyLeuAsn 76
Db 1477 GTTACCATCGAAGCTCAAGCCGCAACGCTACTTTAAAGCAGCGCTGCGATCAGT 1536
QY 77 AlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspAspGlnSerProIleSer 96
Db 1537 GTCACACCTACTCAATATACGTTGATGCTAAGAGTGGCAATGTTACCGCCCAACTTAC 1596
QY 97 ArgIleGlyGluGlnSerProProLeuGlyLeuAsp-----MetSerVal 111
Db 1597 ACATTTGGCGTGAARACCCAGCAGCTTAACAGTGTGGCAGCTAGTCAATAAATTTAGTGT 1656
QY 112 IleGluGluThrThrProLeuSerLeu-----GluGluLeuPheAlaGlnGluSerThr 129
Db 1657 AAGGGTAGTGGTACGAACAATAGCTTAGTTACCGCGCAACATTTGGCAGCTATCTAAT 1716
QY 130 GluMetGlyIleAsnProAsnAsp-----TyrIleProGluTyrGlnGly 144
Db 1717 GAAGTCAATCGAACCGCTGCTACAGTCTCTACAAAGCTTTACCGTTAAAGAGAACGACGAT 1776
QY 145 GluGlnProAsnSerGluValValValProThrThrLeuGluProGluLysProGlyLeu 164
Db 1777 GATGCGCCCAAGCTATCACCGTGGCTAAAGATACGACAAA- 1818
QY 165 IleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAlaLysValProArgLeuLys 184
Db 1819 -----AATGCGCGCAGTCAGCATCTTAAACCTCAA 1851
QY 185 AlaLys-----PheTyrGlnSer 190
Db 1852 GGTAAACAGGTCTAACCGTTGCTACCAAAAGAGTGTACCGTTACCTTTGGGCTTAC 1911
QY 191 SerGlnSerGlyGluThrSerAlaIleGlySerHisGlnLysThrGluProTyrAla 210
Db 1912 CAAGATAGCGTCTGACG-----ATTGGCAAAAGCACCCTTAACACGATGGC---TTG 1962
QY 211 AsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAla-----Met 225
Db 1963 ACTGTTAAGATACCAAGACAAATC---CAAGTCGGTGTCTAATGGCATTTAAATTTACT 2019
QY 226 AspLeuAsnGlySerIlePro-----ArgLeuArgGlnThrAlaLeuValAlaAlaArg 243
Db 2020 AATGTGAATGTAGTAAATCCAGGTACTGGCAATGCAATACCGTCGCATTACCGAGAT 2079
QY 244 AlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAsp 263
Db 2080 AAAATTGGCTTTGTGTTCTGTATGGTGCAGTT----- 2112
QY 264 ValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluVal- 282
Db 2113 -----GATCAACAAACCTT---TATCTTGATCAAGCAAGCTACAAAGTTGGC 2157
QY 283 -----ArgGlyGluGlyAlaAlaAspAspLysAlaPheThrThrValAla 296
Db 2158 AATGTTAAGATTACCAACACTGGCATTAACGCGTGTAAAGCCATCACAGGCTGTCC 2217
QY 297 AspGluValProLeuLeu----- 302
Db 2218 CCAACACTGGCTAGCATTTGCCGATCAAAAGTAGCGGCAACATAGAACTGGGCAATACAATC 2277
QY 303 -----IleGlyAspValPheHisHisGly---Lys 311
Db 2278 CAAGACAAAGCAAAATCCACGCTGCCAGCATTAATGATATATAATACAGCGCTTTAAC 2337
QY 312 TyrGluThrLysLysAlaLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGly 331
Db 2338 CTAATAAATAATAACACCCCAATTGCTTGTCTCC----- 2373

QY 332 ArgTrpLeuAspArgSerValAspValIle-----LeuProAspAsnThrAlaAspVal 349
Db 2374 -----ACTTATGACATTTGTGACTTTGCCAATGGCAATGCCACCGCCG 2418
QY 350 SerLeuIleTyrAspThrGlnTyrArgPheAspGluValValPhe---PheThr 368
Db 2419 ACAGTAACCCATCATACCGCTAAC-----AAACACAGTAAAGTGGTATATGATGATGAAT 2472
QY 369 IleAspProLysThrAsnGlnLeu-----ThrThrAspProAspLysLeuProValLys 386
Db 2473 GTGGATGATACACCATTCATCTAACGCGCTGATGACAAATAAAACCTTGGCGCTCAA 2532
QY 387 ArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAla 406
Db 2533 ACCACCAAACTGAAC----- 2547
QY 407 ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyr---PheAsnMetValAsnThr 425
Db 2548 ---AAACAAGTGTCTAATGGTAATACAGCAACTAATCTTAATGTTAACTAGTATGATGAA 2604
QY 426 GluIleValPheProGluArgGluGlnIleGlnAsn---AspGlnValSerPheGluGln 444
Db 2605 GATGCGCTTGTTAAGCCCAAGACATCCCGCAAAATCTTAACACCTAGCCAGGAAAT 2664
QY 445 SerSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIle 464
Db 2665 CACACCCAAAGGACAGCAGACACCGCCCTTACAAACCTTTACCGTTAAAGAGCTAGAT 2724
QY 465 GluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAla 484
Db 2725 GAAATAATAATGCTGATGACGCCAACCCCATCACCGTGGGT-----CAAAAGCAACCA 2778
QY 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
Db 2779 AATAAT-----CAAGTCAACACCTCAACACTCAAAAGGTGAACCGTCTTAAT 2826
QY 505 MetProAspAspArg-----ValLeuAlaIleAsnHisAspAspGlyValAsn 520
Db 2827 ATTAACCCGACAAATAATGTTACGTTACCTTTGGCATTAACACCCACCAAGCGTCTTAAA 2886
QY 521 ArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuPro 540
Db 2887 -----GCCGGCAAAAGCACCCTTAAC 2907
QY 541 AspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLys 560
Db 2908 GACGCTGCG-----TTGCTCTATTAATAAACCCACTGCTAGCAACAATAATCAA 2955
QY 561 ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAsp 580
Db 2956 GTCGCTGCTGATGCGTGAAGTTGCCAAGTT----- 2988
QY 581 LysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuVal 600
Db 2989 ---AATAATAATGCTGTGTAGTGTGCTGCG-----ATTGATGGCACAACTCGCAT 3036
QY 601 ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArg 620
Db 3037 ACCAGATCAAAATGGCTTTACTGGCACTAATGCG----- 3072
QY 621 LeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeu 640
Db 3073 ---TCACTTCATAAA-----AGCAAAACCC----- 3093
QY 641 AsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPheGlyHis---SerThr 659
Db 3094 -----CACCTAAGCAAA 3105
QY 660 AsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGln 679
Db 3106 GACCGCATTAACCGCGTGGTAAAAAGATACCAACATTCAGTCAGGTGAGATTGCCCAA 3165
QY 680 Asn-----GlyGlyTrpAsnArgThrTyrSerLeuArgTyrArgLeu 693

Db	2670	ACTGTTAAAGATACCAACGAACAATC---	CAAGTCGGTGCTAATGGCATTAAATTTACT	2720
Qy	226	AspLeuAsnGlySerIlePro-----	ArgLeuArgGlnThrAlaLeuValAlaAlaArg	243
Db	2727	AATGTGAATGGTAGTAATCCAGGTACTGCAATG	CGCAATACCGCTCGCATTTACCAGAGAT	2786
Qy	244	AlaValGlyTyrTyrAspIleAspLeuSerIle	lleAlaArgAsnSerIleGlyGluValAsp	263
Db	2787	AAAATTGGCTTCTGCTGTTCTGATGTGCAGTT	-----	2819
Qy	264	ValIleIleHisAspLeuGlyGluProValTyr	IleAspTyrArgAlaValGluVal---	282
Db	2820	-----GATCAACAACAACT---	TATCTTGATCAAGACAAGCTACAGCTTGGC	2864
Qy	283	-----ArgGlyGluGlyAlaAspAspLysAla	PeHeThrThrValAla	296
Db	2865	AATGTTAAGATTACCAACTGGCATTAACGCAG	TGGTAAAGCCATCAGGGCTGCTCC	2924
Qy	297	AspGluValProLeuLeu-----	-----	302
Db	2925	CCAACACTGCCTAGCATTCGGATCAAGTAGCC	CAACATAGAACTGGGCAATACAATC	2984
Qy	303	-----IleGlyAspValPheHisHisGly--	--Lys	311
Db	2985	CAGACAAAGACAANAATCCAAACGCTGCCAG	CAPTAATATATTAAATACAGCGTTTAAC	3044
Qy	312	TyrGluThrLysLysAsnLeuIleGluAsnAla	SerAlaGluHisGlyTyrPheAspGly	331
Db	3045	CTAAAAATAATCAACCCCATTCGCTTCTCTCC	-----	3080
Qy	332	ArgTrpLeuAspArgSerValaspValIle---	-----LeuProaspAsnThrAlaaspVal	349
Db	3081	-----ACTTATGACATGTGTACTTTGCGCAAT	TGGCAATGGCAATGCCACCACCGCC	3125
Qy	350	SerLeuIleTyrAspThrGlyThrGlnTyrArg	PheAspGluValValPhe---PheThr	368
Db	3126	ACAGTAACCCATGATACCCGCTAAC-----	AAACCACTAAAGTGGTATATGATGTGAAT	3179
Qy	369	IleAspProLysThrAsnGlnLeu-----	ThrThrAspProaspLysLeuProValLys	386
Db	3180	GTGGATGATACAACCATTCATCTAACAGGCAC	TGATGACATAAAAACTTGGCGTCAAA	3239
Qy	387	ArgGluLeuLeuGlnLeuLeuThrValAsnMet	GlyGluAlaTyrAsnLeuGlnAla	406
Db	3240	ACCACCAAACTGAAAC-----	-----	3254
Qy	407	ValArgAlaLeuSerAsnAspLeuIleAlaThr	ArgTyr---PheAsnMetValAsnThr	425
Db	3255	---AAAAACAAGTCTAATGGTAAATACAGCACT	AACTAATTAATGTTAACTCTAGTGTGAAT	3311
Qy	426	GluIleValPheProGluArgGluGlnIleGln	Asn---AspGlnValSerPheGluGln	444
Db	3312	GATGCCCTTGTAAACCCCAAGACATCGCCGAA	NAATCTTAACACCCCTAGCCCAAGGAATT	3371
Qy	445	SerSerSerSerArgThrGluProAlaGlnVal	AspGluSerThrLeuGluProValIle	464
Db	3372	CACACCACCAAGGCACAGCACCGCCCTACAA	CCCTTTTACCGTTTAAAAAGGTAGAT	3431
Qy	465	GluThrValGluLeuThrAspGlyIleLeuMet	aspIleSerProIleGluPheSerAla	484
Db	3432	GAAATAATAATGCTGATGACGCCAACGCCAT	CAACCGTGGGT-----CAAAAGAAGCGCA	3485
Qy	485	SerAsnLeuIleGlnAspLysLeuAsnLeuVal	AlaAlaLysAlaArgHisLeuTyrAsp	504
Db	3486	ATAAAT-----CAAGTCAACCCCTAACACT	CAAGGTGAAACGGGTCTTTAAT	3533
Qy	505	MetProaspAspArg-----	ValLeuAlaIleAsnHisaspGlyValAsn	520
Db	3534	ATTAACCAACCCGACAAAAAATGGTACGGTT	ACCTTTTGGCATTAACACCAACGCGGTCTTAA	3593
Qy	521	ArgSerIleLeuGlyArgIleSerAspAlaVal	SerAlaValAlaArgAlaIleLeuPro	540
Db	3594	-----GCCGCAAAAGACCCCTTAAC	3614	

Qy	485	SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp	504
Db	3486	ATAATA-----CAAGTCAACCCCTAAACACCTCAAAAGGTGAAACCGTCTTAAT	3533
Qy	505	MetProAspAspArg-----ValLeuAlaIleAsnHisAspAspGlyValAsn	520
Db	3534	ATTAAAAACCACAAAATAAGTACGGTTACCTTTGGCATTTAAACACCAACGCGTCTTAAA	3593
Qy	521	ArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaAlaIleLeuPro	540
Db	3594	-----GCGCGCAAAAGCACCTTAAC	3614
Qy	541	AspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLys	560
Db	3615	GACGGTGGC-----TTGTCTATTAAAAACCCCACTGGTAGGCAACAAATCCAA	3662
Qy	561	ThrProIaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAsp	580
Db	3663	GTCGGTCTGATCGCGTGAAGTTGCCAAGTT-----	3695
Qy	581	LysProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuVal	600
Db	3696	---AATAATAATGGCTGTGTAGGCTGGC-----ATTGATGGCAACAACCTCGCATT	3743
Qy	601	ThrLysPheGluHisAsnLeuLeuAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArg	620
Db	3744	ACCAAGATGAAATGGCTTTACTGGGACTAATGGC-----	3779
Qy	621	LeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeu	640
Db	3780	---TCACCTTGATAAA-----AGCAAAACCC-----	3800
Qy	641	AsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHis---SerThr	659
Db	3801	-----CACCTAAGCAAA	3812
Qy	660	AsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGln	679
Db	3813	GACGGCATTAACGCAGGTGTAAAGAGATTACCAACATTCATCAGGTGAGATTGCCCAA	3872
Qy	680	Asn-----GlyGlyTyrAsnArgThrTyrSerLeuArgTyrArgLeu	693
Db	3873	AACAGCCATGATCCTGTGACAGCGCGC-----AAGATTTATGATTTAAAAACCGAACTT	3926
Qy	694	AspLysLeuLysThrGlnAlaProProGluThrTyrGlnAspLeuProValAspPheVal	713
Db	3927	GAA-----	3929
Qy	714	AsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys---ThrVal	732
Db	3930	---AACAAAATCAGCAGTACTGCCAAAACAGCACAAAACCTCATTTACAGAAATTCACAGTA	3986
Qy	733	AlaAsp-----AsnLeuValAsnProMetArgGlyTyrArgGlnArg	746
Db	3987	GCAGATGAACAAGGTAATAACTTTACGGTTAGTAAACCCCTTACTCCAGTTATGACACCTCA	4046
Qy	747	TyrSerLeuGluVal-----GlySerSerGlyLeuValSerAspAlaAsnMet	762
Db	4047	AAGACCTCTGATCTCATCACCCTTCAGGTGNAACGGCATTTACCAACAGGTTAAATAAAA	4106
Qy	763	AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer	782
Db	4107	GGTGTGGTGGCTGTGGCATT-----GACCAAAACCAAGGCTTA	4145
Qy	783	AsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrPheSerAspAsnPhe	802
Db	4146	ACCACGCCCTAAGCTGACCGTGGGTAAATAATGCAAGGATTCATGTACAGACGCAA	4205
Qy	803	AsnHisValProTyrArgLeuArgPheAlaGlyGlyAspGlnSerIleArgGlyTyr	822
Db	4206	AAT-----GGTCAAAATACCATCACAGGACTA	4232
Qy	823	AlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlnValLeu	842

Db	4233	AGCAACTCTAGCTAGTAAATGTTACCAATGATAAGGT-----	4268
Qy	843	AlaValGlyThrAlaGluTyraAsnTyrGluPheMetLysAsp-----	859
Db	4269	AGCGTAGCGCACACAGACAGGCG---AATAATAATCAAGACGAACAAACCGTGCC	4325
Qy	860	AlaValPheGlyAspLe-----	869
Db	4326	GCCAGCATGTTGATGCTGTAAGCGCAGCGCTTAACTTGAAGCAATGGTGAAGCGGTT	4385
Qy	870	AspLysGlyPheThrAsnAspThr	877
Db	4386	GACTTTGCTCTCCACTTATGACACC	4409
RESULT 27			
US-08-621-944A-1			
: Sequence 1, Application US/08621944A			
: Patent No. 6440425			
: GENERAL INFORMATION:			
: APPLICANT: SASAKI, Ken			
: APPLICANT: HARNNESS, Robin E.			
: APPLICANT: LOOMORE, Sheena M.			
: APPLICANT: KLEIN, Michel H.			
: TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER			
: TITLE OF INVENTION: MEMBRANE PROTEIN OF MORAXELLA			
: NUMBER OF SEQUENCES: 10			
: CORRESPONDENCE ADDRESS:			
: ADDRESSEE: Sim & McBurney			
: STREET: Suite 701, 330 University Avenue			
: CITY: Toronto			
: STATE: Ontario			
: COUNTRY: Canada			
: ZIP: M5G 1R7			
: COMPUTER READABLE FORM:			
: MEDIUM TYPE: Floppy disk			
: COMPUTER: IBM PC compatible			
: OPERATING SYSTEM: PC-DOS/MS-DOS			
: SOFTWARE: PatentIn Release #1.0, Version #1.30			
: CURRENT APPLICATION DATA:			
: APPLICATION NUMBER: US/08/621,944A			
: FILING DATE: 26-MAR-1996			
: CLASSIFICATION:			
: PRIOR APPLICATION DATA:			
: APPLICATION NUMBER: US 08/478,370			
: FILING DATE: 07-JUN-1995			
: CLASSIFICATION:			
: ATTORNEY/AGENT INFORMATION:			
: NAME: Stewart, Michael I			
: REGISTRATION NUMBER: 24,973			
: REFERENCE/DOCKET NUMBER: 1038-587			
: TELECOMMUNICATION INFORMATION:			
: TELEPHONE: (416) 595-1155			
: TELEFAX: (416) 595-1163			
: INFORMATION FOR SEQ ID NO: 1:			
: SEQUENCE CHARACTERISTICS:			
: LENGTH: 6973 base pairs			
: TYPE: nucleic acid			
: STRANDEDNESS: single			
: TOPOLOGY: linear			
: MOLECULE TYPE: DNA (genomic)			
: FEATURE:			
: NAME/KEY: CDS			
: LOCATION: 708...6883			
US-08-621-944A-1			
Alignment Scores:			
Pred. No.: 0.000858			
Score: 136.00			
Percent Similarity: 32.95%			
Best Local Similarity: 19.73%			
Query Match: 2.88%			
DB: 4			

Db	2985	CAAGACAAAGACAAATCCACGCTGCCAGCATTAATGATATATTAATTAATACAGGCTTTAAAC	3044
Qy	312	TyrGluThrLysLysAsnLeuLeuGluAsnAlaSerAlaGluHisGlyTyrPheAspGly	331
Db	3045	CTAAATAATAACAACCCATTGACTTGTCTCC	3080
Qy	332	ArgTrpLeuAspArgSerValaspValle-----LeuProAspAsnThrAlaaspVal	349
Db	3081	-----ACTTATGACATTGTGTACTTTGGCAATGGCAATGCCACCAACCGCC	3125
Qy	350	SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValphe---PheThr	368
Db	3126	ACGTAACCCATTGATACCGCTAAC-----AAAACAGTAAGTGTATATGATGTGAAT	3179
Qy	369	IleAspProLysThrAsnGlnLeu-----ThrThrAspProAspLysLeuProValLys	386
Db	3180	GTGGATGATACACCATTCATCTACAGGCACCTGATGACATAAATAAACTTTGGCGTCNA	3239
Qy	387	ArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyAlaTyrAsnLeuGlnAla	400
Db	3240	ACCACCAAACTGAAC-----	3254
Qy	407	ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyr---PheAsnMetValasnThr	425
Db	3255	---AAAACAAGTCTAATGGTATACAGCACTAACTTAATGTTAATCTAGTAGTAA	3311
Qy	426	GluIleValPheProGluArgGluGlnIleGlnAsn---AspGlnValSerPheGluGln	444
Db	3312	GATGCCCTTGTAAACGCCAAAGACATCGCGGAAATCTTAACACCTTACCAAGAAATTT	3371
Qy	445	SerSerSerArgThrGluProAlaGlnValaspGluSerThrLeuGluProValle	464
Db	3372	CACACCACCAAGGCACAGCAGACCGCCCTACAAACCTTTTACCGTTTAAAAAGGTAGAT	3431
Qy	465	GluThrValGluLeuThrAspGlyIleLeuMetaspIleSerProIleGluPheSerAla	484
Db	3432	GAATAATAATCTGATGACGCCCAACGCCATCATCCGCTGGT-----CAAAGAAGCGCA	3485
Qy	485	SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaLysAlaArgHisLeuTyrAsp	504
Db	3486	ATAAT-----CAAGTCACACCCCTAACACTCAAGGTGAACCGTCTTAAT	3533
Qy	505	MetProaspArg-----ValLeuAlaIleAsnHisAspAspGlyValasn	520
Db	3534	ATTAACCAACCAAAATGTCGGTTACCTTTGGCATTAACACCAACGCGGTCTTAAA	3593
Qy	521	ArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuPro	540
Db	3594	-----GCGGCAAAAGCAACCCCTAACAC	3614
Qy	541	AspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLys	560
Db	3615	GACGGTGGC-----TTGTCTATTAAACCCCACTGGTAGCGCAACAATCCAA	3662
Qy	561	ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAsp	580
Db	3663	GTCGGTCTGATCGGTGAAGTTTGCAAGTT-----	3695
Qy	581	LysProArgaspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuVal	600
Db	3696	---AATAATAAGTGTGTAGTGGTGGC-----ATTGATGGCACAACCTGCAT	3743
Qy	601	ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArg	620
Db	3744	ACCAGATGAATGGCTTACTGGGACTAATGGC-----	3779
Qy	621	LeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeu	640
Db	3780	---TCACCTGATAAA-----AGCAACACC-----	3800
Qy	641	AsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHis---SerThr	659
Db	3801	-----CACTTAAGCAAA-----CACTTAAGCAAA	3812

D	b	3615	GACGGTGC-----TTGCTATATAAAACCCCTGGTAGGCAACAATCCAA	3662
Q	y	561	ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAsp	580
D	b	3663	GTGGTGTCTGATGGCGTGAAGTTTGCAAGTTT-----	3695
Q	y	581	LysProArgAspGLYGLInIleGlyLeuGlyTTPGlySerAspThrGlyThrArgLeuVal	600
D	b	3696	---AATAAATAAGTGTGTAGGTGCTGC-----ATTGATGGCACAACTCGCATT	3743
Q	y	601	ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArg	620
D	b	3744	ACCAAGATCAAAATTGGCTTTACTGGGACTAATGCG-----	3779
Q	y	621	LeuSerLysAlaAspLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeu	640
D	b	3780	---TCACTTCATAAA-----ACCAACCC-----	3800
Q	y	641	AsnAspGlnLeuArgAlaThrLysGlyTyrGlnGlnGluValPheGlyHis---SerThr	655
D	b	3801	-----CACCTAAGCAAA	3812
Q	y	660	AsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIleIleGln	679
D	b	3813	GACGGCATTAACACGAGTGGTAAAAAGATTACCAACATTCAAATCAGGTGAGATTGCCCAA	3872
Q	y	680	Asn-----GlyGlyTyrAsnAlaGlyThrTyrSerLeuArgTyrArgLeu	693
D	b	3873	AACAGCCATCATCTGTGACAGCGCG-----AGATTATGATTTAAAAACCGCACTT	3926
Q	y	694	AspLysLeuLysThrGlnAlaProProGluThrTyrGlnAspLeuProValaspPheVal	713
D	b	3927	GAA-----	3929
Q	y	714	AsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys---ThrVal	732
D	b	3930	---AACAAATCAGCAGTACTGCCAAACAGCACAAAACCTCATACAGCAATTCACAGTA	3986
Q	y	733	AlaAsp-----AsnLeuValAsnProMetArgGlyTyrArgGlnArg	746
D	b	3987	GCAGATCAACAAGTAATAACTTTACGGTTAGTAAACCCCTTACTCCAGTTATGACACCTCA	4046
Q	y	747	TyrSerLeuGluVal-----GlySerSerGlyLeuValSerAspAlaAsnMet	762
D	b	4047	AAGACCTCTGATCATCACCTTTGCAGGTGAACCGCATTTACCACAAAGGTAAATAAA	4106
Q	y	763	AlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySer	782
D	b	4107	GGTGTGTGCGTGTGGGCATT-----GACCAACCAAGGCTTA	4145
Q	y	783	AsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleThrSerAspAsnPhe	802
D	b	4146	ACCACGCTAAGCTGACCGGTGGTAATAATAATGCAAGGCATTTGTCATTACACAGCCAA	4205
Q	y	803	AsnHisValProTyrArgLeuArgPhePheAlaGlyAspGlnSerIleArgGlyTyr	822
D	b	4206	AAT-----GGTCAAAATACCATCACAGGACTA	4232
Q	y	823	AlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeu	842
D	b	4233	AGCAACACTCTAGCTAATGTTACCAATGATAAGGT-----	4268
Q	y	843	AlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAsp-----LeuArgLeu	859
D	b	4269	AGCGTAGCCACCAACAGAACAGGCG---AATAATAATCAAGACGAACAAACCCGTGCC	4325
Q	y	860	AlaValPheGlyAspIle-----GlyAsnAlaTyr	869
D	b	4326	GCCAGCATTTGTTGATGTGCTAAGCCGAGCGCTTAACTTGCCAAGCAATGGTCAAGCGGTT	4385
Q	y	870	AspLysGlyPheThrAsnAspThr	877
D	b	4386	GACTTTGTCTCCACTTATGACCC	4409

RESULT 29

US-08-431-718C-1
; Sequence 1, Application US/08431718C
; Patent No. 6335018
; GENERAL INFORMATION:
; APPLICANT: Ken, Sasaki
; APPLICANT: Robin, Harkness E.
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT MAJOR OUTER MEMBRANE PROTEIN OF
; FILE REFERENCE: 1038-429 MIS
; CURRENT APPLICATION NUMBER: US/08/431.718C
; CURRENT FILING DATE: 1995-05-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6975
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
US-08-431-718C-1

Alignment Scores:
Pred. No.: 0.000859 Length: 6975
Score: 136.00 Matches: 191
Percent Similarity: 32.95% Conservative: 128
Best Local Similarity: 19.73% Mismatches: 351
Query Match: 2.88% Indels: 298
DB: 4 Gaps: 49

US-09-914-168-2 (1-919) x US-08-431-718C-1 (1-6975)

QY 22 ProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHis 41
DB 2064 CCATATTGGGATAAAACAACTTAAAGTGGGTAGTGTGCAATTACCATAGACAATGGC 2123

QY 42 ValProAlaHisAspThrAlaIle-----ArgLeuAsnAla-----AlaGlyLeuAsn 76
DB 2124 ATTGATCGAGGTATAAAAGATCAGTAATCTTCCAAAGGTACAGTGTAAAGATGGC 2183

QY 61 LeuThrProGluGlnIleGlnAla-----ArgLeuAsnAla-----AlaGlyLeuAsn 76
DB 2184 GTTACCATCGAACAGCTCAAGCGCCAGCCTACTTTAAACGCGAGCGTGGCAGT 2243

QY 77 AlalysProGlnSerGlnAlaLeuAspValValAsnPheAspAspGlnSerProIleSer 96
DB 2244 GTCACCTACTGAATATCAGTTGTGCTAAGAGTGGCAATGTACCGCCCAACTTAC 2303

QY 97 ArgIleGlyGluGlnSerProProLeuGlyLeuAsp-----MetSerVal 111
DB 2304 AACATTGGCGTGAACACCCGAGCTTAACAGTGTACAGTGTAGTATAATTTAGTGT 2363

QY 112 IleGluGluThrThrProLeuSerLeu-----GluGluLeuPheAlaGlnGluSerThr 129
DB 2364 AAGGTAGTGTGAGCAACAATAGCTTAGTTACCGCCGCAACATTTGGCAAGCTATCTAAAT 2423

QY 130 GluMetGlyIleAsnProAsnAsp-----TyrIleProGluTyrGlnGly 144
DB 2424 GAAGTCAATCGAAGCGGTGACAGTGTCTACAAAGCTTTACCGTTAAAGAAAGAGAGAT 2483

QY 145 GluGlnProAsnSerGluValValValProProThrLeuGluProGluTysProGlyLeu 164
DB 2484 GATGACGCCAACGCTATCACCCTGGTGAAGATACGACAAA----- 2525

QY 165 IleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLys 184
DB 2526 -----AATGCCGCGCAGTACGATCTTAAAACTCAAA 2558

QY 185 AlaLys-----PheTyrGlnSer 190
DB 2559 GGTAAAAAGCGTCTAAGGTTGCTACCAAAAAAGATGGTACGTTTACCTTTGGCTTACG 2618

QY 191 SerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAla 210
DB 2619 CAAGATACGGGTCTGACC-----ATTGGCAAAAGCACCCCTAAACACAGATGGC---TTG 2669

QY 211 AsnIleLysAlaAlaLeuLeuAspIleThrGlnGluSerAla-----Met 225
DB 2670 ACTGTTAAAGATACCAACGACAAATC---CAAGTCGGTGTCTAATGGCATTTAAATTTACT 2726

QY 226 AspLeuAsnGlySerIlePro-----ArgLeuArgGlnThrAlaLeuValAlaAlaArg 243
DB 2727 AATGTGAATGCTAGTAATCCAGTACTGGCATTTGCAATACCGCTGCGCATTTCCAGAGAT 2786

QY 244 AlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAsp 263
DB 2787 AAAATTGGCTTTGGTGTCTGTGTCAGTT----- 2819

QY 264 ValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluVal--- 282
DB 2820 -----GATACAAACAACT---TATCTTGATCAAGACAAAGCTACAGTTGGC 2864

QY 283 -----ArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAla 296
DB 2865 AATGTTAAGATTACCAACACTGGCATTAACGCGGTGTAAGCCATCACAGGCGCTGTC 2924

QY 297 AspGluValProLeuLeu----- 302
DB 2925 CCAACACTGCGCTAGCATTTGCGGATCAAAGTAGCGCAACATAGAACTGGCAATACAACT 2984

QY 303 -----IleGlyAspValPheHisHisGly---Lys 311
DB 2985 CAGACAAAGACAAATCCAACGCTGCCAGCATTAATGATATATATAATACAGGCTTTTAA 3044

QY 312 TyrGluThrIlyLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGly 331
DB 3045 CTAATAAATAATAACAACCCCATTTGCTCTCC----- 3080

QY 332 ArgTrpLeuAspArgSerValAspValle-----LeuProAspAsnThrAlaAspVal 349
DB 3081 -----ACTTATGACATTTGTTGACTTTGCCAATGCCAATGCCACCGCC 3125

QY 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhe---PheThr 368
DB 3126 ACAGTAACCATCATATACCGCTAAC-----AAACACGATAAGTGGTATATGATGTGAAT 3179

QY 369 IleAspProLysThrAsnGlnLeu-----ThrThrAspProAspLysLeuProValLys 386
DB 3180 GTGGATGATACAACCATTCATCTAACAGGCACCTGATCACAAATAAAACTTGGCGTCAA 3239

QY 387 ArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAla 406
DB 3240 ACCACCAACTGAAC----- 3254

QY 407 ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyr---PheAsnMetValAsnThr 425
DB 3255 ---AAAACAGTGCTAATGTGTAATACAGCACTAACTTTAATGTTAACTCTAGTGATGA 3311

QY 426 GluIleValPheProGluArgGluGlnIleGlnAsn---AspGlnValSerPheGluGln 444
DB 3312 GATGCGCTTTGTTAAACGCCAAAGACATCGCGGAAATCTAAACACCCCTAGCAAGGAAT 3371

QY 445 SerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIle 464
DB 3372 CACACCCCAAGGCACACAGCACCGCCCTTACAAACCTTTACCGTTTAAAGGTAGAT 3431

QY 465 GluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAla 484
DB 3432 GAAATAATAATGCTGATCAGCCCAACGCTACCGTGGT-----CAAAAGAACGCA 3485

QY 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
DB 3486 AATAAT-----CAAGTCAACCCCTAACACATCAAAAGGTCAAAAGGTCTCTTAAT 3533

QY 505 MetProAspAspArg-----ValLeuAlaIleAsnHisAspGlyValAsn 520

QY 81 -----SerGlnAlaLeuAspValValAsnPhaAspGlnSerProle 95
Db 461819 GCGGCCATCACCGCGCGCAGAGCTG---GCCACCGAGCTGGACGACGAG--- 461772

QY 96 SerArgIleGlyGluGlnSerProLeuGlyLeu-----AspMetSerValle 112
Db 461771 ---TAGCTCTCCAGGACAGCTGATGTCGGGCTGGCCACCGGTGACTCCGAGCTGCC 461715

QY 113 GluGluThr-----ProLeuSerLeuGluLeuLeuPheAlaGlnGlu 127
Db 461714 AAGCTGTTCACGGCGCAGCGCGCTCGCGCAGCGCTGCGGAGGCGCTTCTGCTCAAGTG 461655

QY 128 SerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnPro 147
Db 461654 CGCGCAGCGCGCGGTACACAGCGCCCGAGACCGGAGCGGACCTATCAGCGCTGCAGAAG 461595

QY 148 AsnSer-----GluValValValProProThrLeu----- 157
Db 461594 TACTCCACCGACCTGACCGCGCGCGCGCGAGGCAACTCGACCGCGTCTCGCGCGC 461535

QY 158 GluProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAsp----- 175
Db 461534 GACAACGAGATCCCGCGCTGGTGCAGGTGCTGCTCCGCTCGCACCAAGAACACCCCGTG 461475

QY 176 -----GlyValAsnLys-----ValProArgLeu 183
Db 461474 CTGATCGGTGAGCGCGCTCGGCAAGACCGCGATCGTGGAGGCGCTGCGCGCTCGATCTCGGC 461415

QY 184 LysAlaLysPheTyrGlnSerGlnSerGlyGluThr----- 196
Db 461414 GTGCGCGCGCAGCTGCGCGAGAGCTTGGCGGACAAAGACCATCTGCTGCGCTCGATCTCGGC 461355

QY 197 SerAlaIleGlySerSerHisGlnLysThrCluProTyrAlaAsnIleLysAlaLeu 216
Db 461354 TCGATGTGCGCGCTCCAAATACCGCGCGAATTCGAGGAACGGCTCAAGGCGTCTCTC 461295

QY 217 GluAspIle----- 219
Db 461294 GACGACATCAAGAACTCGCGCGCGCAATCATCATCGTTTCATCGAGGCTGCACACCATC 461235

QY 220 -----ThrGlnGluSerAlaMetAspLeuAsnGlySerIle---ProArg 233
Db 461234 GTCGGCGCGCGCGCCACCGCGGAGGCGGCGATGACCGCGCAACATGATCAAGCGCATG 461175

QY 234 LeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSer 253
Db 461174 CTGCGCGCGCGCGGTTA-----CGGCTGTCGGGCGCACCGCTGGAC----- 461130

QY 254 IleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGluProVal 273
Db 461129 GAATACCGCAAGCACATC---GAGAAGGACGCGCGCTCGAGCGCGCTTTCACACAGTG 461073

QY 274 TyrIleAspTyrArgAlaValGluValArgGlyCluGlyAlaAspAspLysAlaPheThr 293
Db 461072 TACGTGCGGCGCGCTGGTGGAG----- 461049

QY 294 ThrValAlaAspGluValProLeuLeuIleGly-----AspValPhe-----HisHis 309
Db 461048 -----CACACCATCGCATCTCTCGCGGGCTCAAGACCGCTCAGAGGTGCACAC 460998

QY 310 GlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPhe 329
Db 460997 GGGTGGCATCACCGACTCGCGCTGGTGGCAGCTGCCACTTGGACGCGCGTATATC 460938

QY 330 AspGlyArgTrpLeuAspArgSerValAspValIleLeuProAsnThrAlaAspVal 349
Db 460937 ACCGCGCGCTTC-----CTGCGCGCAACAGCGCATCGAC--- 460905

QY 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspCluValPhePheThrIle 369
Db 460904 ---CTGTCGACGAGGCGCGCGCTGCGGATGGAG-----ATC 460866

QY 370 AspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgCluLeu 389
Db 460865 GACTCGCGCGCGCTGAGATC---GACGAGTCCAGCGGCTGGTCCCGGCTCGAGATC 460809

QY 390 LeuGluGlnLeuLeuThrValAsnMetClyGluAlaTyrAsnLeuGlnAlaValArgAla 409
Db 460808 GAAGAGATGCGCTGTCCAAAGAAGACGAGCGC---TCGCGCGAGCGGTTCGCCAAG 460752

QY 410 LeuSerAsnAspLeu-----lleAlaThrArgTyrPhe 420
Db 460751 CTGCGCTCCGAGCTGGCGGACCAAGAGAGTTGGCCGAGCTCACACCGCTGGCAG 460692

QY 421 AsnMetValAsnThr-----GluIleValPheProGluArgGluGlnIleGlnAsnAspGln 439
Db 460691 AACGAGAAGAAGCGATCGAAATCGTCGCGACCTCAAGGAGCAGCTGGAA----- 460641

QY 440 ValSerPheGluGlnSerSerSerSerArgThrGluProAlaGlnValAspGluSerThr 459
Db 460640 -----GCCCTGCGCGGGGAATCCGAGCGGCGCAA----- 460611

QY 460 LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro 479
Db 460610 -----CGCGACGCGCACCTGGCCCAAGCGCGCGAG 460581

QY 480 IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsn-----LeuValAlaAla 497
Db 460580 CTGCGCTACGAGCGCATCCCGCGAGGTGGAGAAAGCTCAGCGCGGTTCGCGCAGGCG 460521

QY 498 LysAlaArgHisLeuTyrAspMet-----ProAspAsp----- 508
Db 460520 CAGGCGCGGAGCAGGTGATGCTCAAGGAGGAGTCCGTCGCCACGACATCCGCGGAGGTG 460461

QY 509 -----ArgValLeu-----Ala 512
Db 460460 GTGTCGCGGTGGACCGGCATCCCGCGCGTCTGCTGCAAGCGAGACCGCAAGCTG 460401

QY 513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532
Db 460400 CTGCGCATGGAAGACGAGCTGGGCAAGCGGTTCATCGGCGAG---AAGCGCGCGGTACC 460344

QY 533 AlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
Db 460343 GCAGTCTCTCATCGGTG-----CGCGCAGCGCGCGCGGTGTCCGAC----- 460299

QY 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValPro 572
Db 460298 -----CCCAACCGCGCCACCGGGCG----- 460278

QY 573 LeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrGly 592
Db 460277 ---TTTATGTTCTCGCGCC-GACCGGTGTGGCAAGACCGAGCTGGCCAAAGCGCTGGC 460222

QY 593 SerAspThrGlyThrArgLeuValThrLys-PheGluHisAsnLeuIleAsnArgAsp-- 611
Db 460221 -----CGACTTCTGTTCGACGACGAGCGGCGGATGTCGCGCATCCACAT 460177

QY 611 ----- 611

Db 460176 GACCGAGTACGCGGAGAGACACACCGTGGCTGCGTTCGCGCGCGCGCGCTATGT 460117

QY 612 -GlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLys-----GlyVally 629
Db 460116 GGGATACGAGGCGCGGTCAGCTACCGAGCGGTGGCGCGCGCTCCMTACACGCTGT 460057

QY 629 sLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuG 649
Db 460056 GCTGTTTCGACGAGATCGAGAAGCGCACCG-----GACGTGTCGAGCTGCTGTGCA 460003

QY 649 yTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLe 669
Db 460002 GGTCTCTCGACGAG-----GGCGCGCTCACCGAGCGGCGCGCGCGCTCGACTTCC 459950

QY 669 uGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArg 685

Db 459949 GCAACACCATCTTGATCTGACGTCCAACTGGGGTGGGGTGGCAGCGC 459901
:::|||| :::: ||||| |||

RESULT 31

US-09-453-702B-243/C

; Sequence 243, Application US/09453702B

; Patent No. 6365723

GENERAL INFORMATION:

APPLICANT: Blattner, Frederick R.

; Burland, Valerie T.

; Perna, Nicole T.

; Plunkett, Guy

; Welch, Rod

TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157

; NUMBER OF SEQUENCES: 265

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Quarles & Brady

; STREET: 1 South Pinckney Street

; CITY: Madison

; STATE: WI

; COUNTRY: US

; ZIP: 53701-2113

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 8.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/453,702B

; FILING DATE: 03-Dec-1999

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/110,955

; FILING DATE: 04-DEC-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Seay, Nicholas J.

; REGISTRATION NUMBER: 27386

; REFERENCE/DOCKET NUMBER: 960296.95017

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (608) 251-5000

; TELEFAX: (608) 251-9166

; INFORMATION FOR SEQ ID NO: 243:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5629

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; SEQUENCE DESCRIPTION: SEQ ID NO: 243:

US-09-453-702B-243

Alignment Scores:

Pred. No.: 0.000761 Length: 5629
Score: 135.00 Matches: 205
Percent Similarity: 31.34% Conservative: 157
Best Local Similarity: 17.75% Mismatches: 365
Query Match: 2.86% Indels: 428
DB: 4 Gaps: 51

US-09-914-168-2 (1-919) x US-09-453-702B-243 (1-5629)

QY 54 AlaGlyAsnProValLeuThrProGluGlnIleGlnAlaArgLeuAsnAlaLa 73
||| ||| ||| :::: ||||| |||
Db 3289 GCCGGAATATCCCGCCTTAAAGCAAGGACTGCACCTAAGGCGACACTTAGCCTTCAC 3230
QY 74 GlyLeuAsnAlaLysProGlnGlnAlaLeuAspValValAsnPheAspGlnSer 93
||||| ||||| ||||| :::: ||||| |||
Db 3229 AACTGATGCTCCACTGCAATCGGCATATTTATACATTTAGAGCAAAACAGATAGGCT 3170
QY 94 ProIleSerArgIleGlyGlnSerProProLeuGlyLeuAspMetSerValIleGlu 113
::: ||| :::: ||| :::: |||
Db 3169 CCGCTTGCACATATATACGCGCAATAATATGACGTTTACGCGCATAAAGACATTTAT 3110

QY 114 GluThrThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGluMetGlyIle 133
||| ||| :::: ||||| |||
Db 3109 ACCCTCAGC-----ATCAACGTCATGATGAGAGTGAT 3077
QY 134 AsnPro---AsnAspTyrIleProGluTyrGlnGlyGluGlnProAsnSerGluValVal 152
||||| :::: ||||| :::: ||||| |||
Db 3076 AATCCCTGACAAATCATCAGTCACCTTTAAGAATCAAAAAGAACGCGGAGTTTGTG 3017
QY 153 ValProPro----- 155
|||||
Db 3016 GAACCCCGCAGCAAAATACGGATGCATATGTTGTCGCCACAATAAACATGTAAGTCAG 2957
QY 156 -----ThrLeuGluProGluLysPro---GlyLeuIleLysArgLeuTyr 169
||||| :::: ||||| |||
Db 2956 GTTTCGGAAGAAATACGATTAGCGCCACGCTGCCAAATGGTTTTTCACACACGGATAATT 2897
QY 170 AlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyrGln 189
||||| :::: ||||| |||
Db 2896 GCGAAATTCGTTAGCGAT-----TCGAGTACGCCA-----AAATTCAAACAA 2855
QY 190 SerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr----- 206
||||| :::: ||||| |||
Db 2854 CTGGTTGCCGATCCAGATACCATTTATGCTGGCAACAGCAGCGGACGACTCTGACCGCC 2795
QY 207 -----GluProTyrAlaAsnIleLys----- 213
||||| :::: ||||| |||
Db 2794 ATCATCACAGACTTTCATAACAACCGTTAAAGATATGAAGATGAATTTGTGGCACCT 2735
QY 214 -----AlaAlaLeuGluAspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIle 231
::: ||||| :::: ||||| |||
Db 2734 GTTGGCTCGCAACTGGACACACGACG-----GCCACACAGACGACGCTCCGCTATTGTG 2681
QY 232 ProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAsp 251
::: ||| :::: ||||| |||
Db 2680 -----CGGTGCGACCTGACCATGCTCAAAAGCT---GGTAGCTATTCCGTCGAT 2636
QY 252 LeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGlu 271
||||| :::: ||||| |||
Db 2635 GCCTCGCTT-----GAGGTGGATAAAATATTACCACGACGCTCGGTCACG 2594
QY 272 ProValTyrIleAspTyrArgAlaValGluVal-----ArgGlyGluGlyVala 287
::: ||| :::: ||||| |||
Db 2593 ATCAGCTGCTCCCAACAGGGAACAAATCGGTAATGACCTTGAATGCCGGTCCGGCAGT 2534
QY 287 ----- 287
Db 2533 GCGATCGCTAACAAATACAAATATCGTTACCTGACTGCCAGTGTGAAGATGTTTATGGA 2474
QY 288 -----AspAspLysAlaPheThrThrValAla----- 296
::: ||| :::: ||||| |||
Db 2473 CACCCGTTGCCGATGAGGATGTGAAATTTACCTTCCAGCCTCCATGACCGGAACTTC 2414
QY 297 -----AspGluValProLeuLeuIle 303
||| ||| ||| |||
Db 2413 AGCTAAGTAGTGAACACCCCGCCGATGCAACAGCGTGATGCCGTGTCACATTGCGGA 2354
QY 304 GlyAspValPheHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSer 323
||| ||| :::: ||||| |||
Db 2353 GGCACA-----AAAGCGGTGAGTTTACAGTTACGCGCGCTGACACAGANAATAACC 2300
QY 324 AlaGluHisGlyTyr-----PheAspGly-----ArgTyrLeu 334
::: ||| :::: ||||| |||
Db 2299 GTTGTCTATCAGCAGTCACTTTTATTGGGGATACAAACAGTCGCGCAGCTCCACGCGCTG 2240
QY 335 AspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeu----- 351
||||| :::: ||||| |||
Db 2239 ACTGCCTCAATTAATTCCTCATTTGTCGGGTAAACAGTACCGGGAGTACCCTGACGGCAACG 2180
QY 352 IleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePhe----- 367
||| ||| :::: ||||| |||
Db 2179 ATCCTGGACGCTTACCANAATCCGCTTAAGACCACTTGGTCTACTTCCAGAGCTAAGCAT 2120
QY 367 ----- 367

Db 2119 GTCACCTCTACCGAAACAGAGAGTCAACCACCAATACGCTGGGTCTACGGCGACCGGTACCAATG 2060
Qy 368 -----ThrIleasp 370
Db 2059 ACCAGCAATATTCGGGACACATAACGCTGCTGGTGACCCGGAAGCGCAAGCTTCGGAT 2000
Qy 371 ProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLys---ArgGluLeu 389
Db 1999 AATAAAACGTTTGTAGTTATACAGTCTACCGGATGAAGTTCGGCGAAGGTAAATAGTATA 1940
Qy 390 LeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGln---AlaValArg 408
Db 1939 ACCGAGCGCAAAACACATACGGTGGGCGGAACACATCACGCTACGGATCTCTCCAG 1880
Qy 409 AlaLeuSerAspLeuIleAlaThrArg 418
Db 1879 GACGGCTTTACAAATGTAATCGCGGTCAACGCGTCAGATTAAAGTGGCGAGCCAACT 1820
Qy 419 -----TyrPheAsnMet 422
Db 1819 AACATTACGATAGCGGATACGGCTTACACCGATAATAACGGTTATGCGGTACGTTAACCTT 1760
Qy 423 ValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPhe 442
Db 1759 CTCAGCACCCAA-----CTGGGGTTTATCAGGTG-----ACGCAACGCTG 1718
Qy 443 GluGlnSerSerSerArgThrGlu 451
Db 1717 GACATAACAGTAGTAGTAAGGTGACGTGAATGTGGCAATGGCAAACTCGAGTTAAACA 1658
Qy 451 ----- 451
Db 1657 TCATCGAAACCAAGAACTACGGTCCATAATAGTAGGGGTATTACGGTGCACCGCAACGGCG 1598
Qy 452 -----ProAlaGlnValAspGluSerThrLeuGluProVal 463
Db 1597 AGAATGCGGGGGTGAATTGATGCCAGGGCAATTTATCACCTTTAGCGTAACGCTGAA 1538
Qy 464 IleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSer 483
Db 1537 GGTGCAACGCTAGCAATACAGGGAAGTCTTACTGACCACTCA-----GGTCAGGCC 1484
Qy 484 AlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyr 503
Db 1483 AAAGTGACGCTGACCACTGACAAAGTGAATGTCTATACCGTTACGGCCATAATGGCCAA 1424
Qy 504 AspMetPro-----AspAspArgValLeuAlaIleAsnHisAsp 516
Db 1423 GATGTTCCGTTGAGCCAGGTAAACGGTTGCGGTTAAGCGCAGATGCTAAACCGGCACAT 1364
Qy 517 -----AspGlyValAsnArgSer--- 522
Db 1363 GTTGTGAGCGTGGCTTCTCTGACACCATCACCGCCGACGCGATCGATAGCAGACC 1304
Qy 523 IleLeuGlyArgIleSerAsp-----AlaValSerAlaValAlaArgAlaIle 538
Db 1303 ATCACTTCACGAGTAGAAGATGATTACGGATTCCCGGTTCAAGGTTCGATATTAGTCAT 1244
Qy 539 LeuProAspGluSerGluAsnGluValIleAspLeuPro----- 551
Db 1243 GGCTTAGCACCAAGCGACCGCGGTAGTTAAATATTCCTCACTACGCGTACCGATCAGTCC 1184
Qy 552 -----GluArgThrAlaLeuAlaAsnArgLys 560
Db 1183 GGGCAAGTCAAGCGGACAAATAACCAAGTACATTGGCAGAAACCTTTAACAGTCATATGTGCAA 1124
Qy 561 ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAsp 580
Db 1123 GTTCTGGCACAGCCAAACCAATCGCAACCAATTACATTGTTGGCGCGCAGCGCGATGAA 1064
Qy 581 LysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuVal 600

Db 1063 -----AGTAAGTCAATT 1052
Qy 601 ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlnLeuArg 620
Db 1051 TTGAATCCGATGTTGACACTCTGAAGCTCAC---TACCAGCAGCGCAAACTTACG 995
Qy 621 LeuSer---GluAspLysGlyValLysLeuTyrAlaThrLysProLeu----- 636
Db 994 CTAACTTGCAGACAACTACGGTAACCGATAGTGAGCTCTGTGATCATCTGCAATTTGTC 935
Qy 637 ---SerHisProLeuAsnAspGlnLeuArg---AlaThrLeuGlyTyrGlnGlnGluVal 654
Db 934 CAGTCAGGCCCTTCGTGAACCTTCTCAAGTTGACGATATTGATTACAGCCCAAGAAAT 875
Qy 655 PheGlyHisSerThr----- 659
Db 874 TATGCGGAGTACACCGTCTACTGCGGAAAGAGGAACAGCGACACTCATTTCCC 815
Qy 660 -----AsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIle 677
Db 814 ATGCTGAACGGGTTTCATCAGGCAAACTTTAAGCATATCGCTGAATCTCATCAATTCGATA 755
Qy 678 IleGlnAsnGlyGlyTrp-----AsnArgThrThrTrpGlnAspLeuProValAspPheVal 693
Db 754 AAAGAAATGTCCCGTCATCTACTGCAAAACCACTACCTTCTCCACGGCT--- 704
Qy 694 AspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheVal 713
Db 703 ---AAATCCCGCAGCGAGGCTTTCGAGGAGGCTATTACACACTC----- 602
Qy 714 AsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAla 733
Db 661 -----AACAAATGATAACTTTGAAGCGGGT-----AAAACCGTTGAT 626
Qy 734 AspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySer 753
Db 625 GATTATATGTTTCAAGTTTCACAGTTCACAGGGTTGG-----GTGCTGTCGATGCT 581
Qy 754 SerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyr 773
Db 580 TCGGGTAAAGTTTCTTTCGCAATATC----- 554
Qy 774 SerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGly----- 791
Db 553 -----GGCGATCAACAGCTCAGTCACAATTAAGCGCTTCCCGCAGCAAGAGGTACAAAC 500
Qy 792 -----IleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyr 807
Db 499 TACCAGACCTTAATTAAGCTCAAGGCTGGTGGTGAATAATGGAAATCATACCAATATC 440
Qy 808 ArgLeuArgPhePheAlaGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeu 827
Db 439 TGGCTA-----GCTGCCAATGCGCTC 419
Qy 828 SerProIleSerAspLysGlyTyr-----LeuThrGlyGlyGln 840
Db 418 TGTCATCTAAAAATGATGATATATCTCCCTGGCATCATACACATTTGACGCTCTGGCGAA 359
Qy 841 ValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAla 860
Db 358 -----AACAAACGCAAGCGGATCA 338
Qy 861 ValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGly 880
Db 337 CTGTATGTGTAATGGGGAACGTT-----GGAGCGTTTTCAGTAATTTCCCAATTTTACA 284
Qy 881 AlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArgValAspValAlaThr 900
Db 283 CCG-----GGAGCTTACTGCAGGT----- 263
Qy 901 GlyValLysGluGluGlyAsnProIleLysLeuHisPhePheIle 915
Db 262 -----GAATCTGATGATTACAGTCGCGCACTACTATGTC 230

RESULT 32

US-09-134-001C-322
; Sequence 322, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 322
; LENGTH: 30549
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-322

Alignment Scores:

Pred. No.: 0, 0146 Length: 30549
Score: 135.00 Matches: 166
Percent Similarity: 32.81% Conservative: 106
Best Local Similarity: 20.02% Mismatches: 295
Query Match: 2.86% Indels: 262
Gaps: 4

US-09-914-168-2 (1-919) x US-09-134-001C-322 (1-30549)

QY 27 GlnAlaLeuAlaGlnGlnAsnProAlaAsn-----IleIleAsnHisValProAlaHis 45
Db 8083 CAAGCAGCTGAAGCAAGCAATACAAATGCGTTATAAATAATGGAGATCCACA 8142
QY 46 AspThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGluGln 65
Db 8143 TCCCAACAAATTAATGATGCTAAA-----AACACAGTTGAACAGGCACAGAGATTAT 8196
QY 66 IleGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAsp 85
Db 8197 GTTGAAGCTAAAGCAACTTAGCTGCTGATAGTACAGTTACAGTTACAAAGC---GCTTATGAT 8253
QY 86 ValValAsnPheAsp-----AspGlnSerPro----- 94
Db 8254 ACGTTAAATAGAGATGTTTAAACAAATGATAAAAGCCAGCATCTGTAAGCGCTATAAT 8313
QY 95 -----IleSerArgIleGlyGlnSerProProLeuGlyLeuAspMetSerValIle 112
Db 8314 GAAGCCATTTCAAATATTAGAAAAGAAATAGATACAGCTAAAGCGGATGCAAGTAGTACT 8373
QY 113 GluGluThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGluMetGly 132
Db 8374 TTGCGAACAACCAANTCCTTCGCTGAACAGTTAGAGACGCTTTAAATAAATAAATACT 8433
QY 133 IleAsnPro-----AsnAspTyrIleProGluTyrGlnGlyGlnProAsnSerGlu 150
Db 8434 GTTCAACCTAAAGTGAATCAAGCAATTCGTTTACTTCAACCAAGAAATAAATTCAGAA 8493
QY 151 ValValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAla 170
Db 8494 CTTGTACAAAGCTAAA-----AAACGTTTA----- 8517
QY 171 ArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAla----- 185
Db 8518 -----CAAGCGCTGTAATGACATACCTCAACACAGAGGTATGACACAACAACA 8568
QY 186 ---LysPheTyrGlnSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGln 204
Db 8569 ATTAATAATATATGACAAACAAGCTGAAGTGAAAGAGCATTACATCTGCCAACAAGA 8628

QY 205 LysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAla 224
Db 8629 GTGATTGAT-----AATGGGGATGCTACAACTCAAGAATATTACTTCTGAAAATCT 8679
QY 225 MetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAla 244
Db 8680 -----AAAGTAGAGCAAGCAATGCAA--GCTTTAACTAATGCTAAA--- 8718
QY 245 ValGlyTyrTyrAspIleAspLeuSerIleIleAtqAsnSerIleGlyGluValAspVal 264
Db 8719 -----AGTATCTGAGA----- 8730
QY 265 IleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGly 284
Db 8730 ----- 8730
QY 285 GluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGly 304
Db 8731 -----GCTGATAAGAATGAGTTACAGACTGCATATAACAAA-----TTAATTGAG 8775
QY 305 AspValPheHisGlyLys-----TyrGluThrLysLysAsn 317
Db 8776 AACGTATCTACCAATGGTAAAAACCGCGAGTATACGTCAATACGAACACCCAAAGCC 8835
QY 318 LeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAsp----- 335
Db 8836 AGATACAAAATCAAATTAATGATGCTAAAAATGAAGCGGAGCAATTTTAGTAAATGAT 8895
QY 336 -----ArgSerValAspValIleLeuProAsp 344
Db 8896 AATCCACAAGTATCACAGTAAGTGAACAAATCAAAGCTATTCAACGCAAAA 8955
QY 345 AsnThrAlaAspValSerLeuIleTyrAsp-----ThrGlyThr 357
Db 8956 TTAACAGAGAGCTATCAACATGCTTCAACACAAGAAATAATACAGAAATAGTCAATGCT 9015
QY 358 GlnTyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThr 377
Db 9016 AAAAAACAGACTTGAATAATGCAGTAAATATACA--GATCCA---ACACACGTATGACT 9069
QY 378 ThrAsp-----ProAspLysLeuProValLysArgGluLeuGlu 391
Db 9070 CNAAGAAACAATTAATAATTACACCGCTAAAAGCGAGAGCTCAAAATGAAATACAAAA 9129
QY 392 GlnLeuLeuThrValAsnMetGlyGluAla----- 401
Db 9130 CGCAACATGATTATTATAATGGAGATGCTACTGCTCAAGATATTCTTCTGAAAAATCT 9189
QY 402 -----TyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeu----- 414
Db 9190 AAAGTAGAGCAAGTATTACAGCATTTACAAAATGCTAAGAATGACTTAAAGAGCTGATAA 9249
QY 415 -----IleAlaThrArgTyrPheAsnMet-----ValAsnThrGluIleValPhe 429
Db 9250 AGAGAATTACAGACTGCATACAAATAACTTATACAAAATGTTAATACCAATCGTAAAAA 9309
QY 430 ProGlu-----ArgGluGlnIleGlnAsnAspGlnValSer 441
Db 9310 CCATCTAGTATTCAAAACTATAAGTCTCAACAGCAAAATATCGAAAAAC----- 9357
QY 442 PheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGlu 461
Db 9358 -----CAATAATATACCGCTAAAAAATGAAGCACATAATGTTCTTTGAAAAATACA---AAC 9408
QY 462 ProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspLysSerProIleGlu 481
Db 9409 CCTACTGTAATGCGAGTACAGATGCTTTACCTAAGATAAATGCAATTCACACAGAGTT 9468
QY 482 PheSerAlaSerAsnLeuIleGlnAspLys----- 491
Db 9469 ACAAAAGCTATTAAATATATTCAAGATAAAGAGATAATAGCGCACTTCTTAGACCAAAA 9528
QY 491 ----- 491

```
Db 9529 GAAAAATTAGATCAAGCGATTAAATAGTCAACCATCACTAAATGGTATGACTCAAGAAATCT 9588
Qy 492 LeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspArgValLeu 511
Db 9589 ATTAATAATTACACACAAAGAGTAGAGACGACAAATATAGCAAGTTCCTGCTGACACT 9648
Qy 512 AlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGly-----ArgIleSer 528
Db 9649 ATTATTAATAATGGGATGGCATCTATTGAACAAATAAGACAGAAATAAATTCGAGTTGAA 9708
Qy 529 AspAlaValSerAlaValAlaArgAla-----IleLeuProAspGluSer----- 543
Db 9709 GAGGCACTAATGACCTTAAGCAAGCAAAACACATTTAAGCGGAGATACAACTTCCTTTA 9768
Qy 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
Db 9769 AAAAGTGAAGTACGGAATTA---AGTAGGAGCGGACACAAACAAACAAAGACCTAGC 9825
Qy 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583
Db 9826 AGTGTT----- 9831
Qy 584 AspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuValThrLysPhe 603
Db 9831 ----- 9831
Qy 604 GluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGlu 623
Db 9832 -----AGTGCCTTATAACAATACTATTCTTCGCTACAACTGAAATTCACACAGACTGAA 9885
Qy 624 AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSer----- 637
Db 9886 AATAGCAATAACTATCATC-----AATAAGCCTATTCTGTTGTAAGAAGTAAAT 9939
Qy 638 -----HisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGln 651
Db 9940 AATGATTGTCATGAAGTAACCAATGAACCAACGCTTAACAGATACAAATTAACCTTATTA 9999
Qy 652 GlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHis 671
Db 10000 CAACCTTTA-----GCGAATAAAGAGCTTAAAGAAAGCTCGTAATCGACTTGAAGT 10053
Qy 672 GluIleSerArgSerIleIleGlnAsnGlyGlyTyrPheAsnArgThrTyrSerLeuArgTyr 691
Db 10054 AAATTAATGAACCC---GTTCAACAGACGCGTATGACTCAACAACTCTGTTGAGAAATAT 10110
Qy 692 ArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAsp 711
Db 10111 AAGCAAGCTAAATAAAGCTCAAAATGAATCTAGTATTGCACAAACTCTT----- 10161
Qy 712 PheValAsnGlyLysProSerGlnGlu 720
Db 10162 ATTAATAATGGTATGATCTGATCAA 10188

RESULT 33
US-09-221-017B-1002/C
; Sequence 1002, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: ROSS, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1002:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...9377
US-09-221-017B-1002

Alignment Scores:
Pred. No.: 0.00211 Length: 9377
Score: 134.50 Matches: 117
Percent Similarity: 34.00% Conservative: 69
Best Local Similarity: 21.39% Mismatches: 188
Query Match: 2.85% Indels: 173
DB: 4 Gaps: 30

US-09-914-168-2 (1-919) x US-09-221-017B-1002 (1-9377)
Qy 507 AspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArg----- 521
Db 2702 GACTTCAGAGTTCAGGAGCTGCCACGATATCGATTTCGCGTGGCGGATACACTGTT 2643
Qy 522 -----SerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIle--- 538
Db 2642 TACTATGGGAGTAGGAGCAGCTATATCCGGCA---TCGGCTCTCACGCGTGGTGTCC 2586
Qy 539 -----LeuProAspGluSerGluAsnGluValIleAspLeu--- 550
Db 2585 GTTACACCGGAGCTTTTTCGCGAGGATGATGTGAACGCTCTTATATCAAGCTGAAT 2526
Qy 551 -----ProGluArg 553
Db 2525 GCCTCCCTATCGTTCGGAACGTAATATCCGATTTGTTGGAGCACAAATGTTAGCATGAG 2466
Qy 554 ThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeu 573
Db 2465 ATTGCTCTGGCGGAT-----AGCTCTCGCCCTTGTGGACTGC 2430
Qy 574 TyrValPheValAlaSerAspLysProArgAspGlyGln-----IleGlyLeuGly 590
```



```

Db 2429 TATATTCTTACCGTTCGGCCCAAGAGCAATCGTTTCGAAGCGCAAGTCTCGGCAACCAAT 2370
      |||...  . ||| ...  :|||
Qy 591 TrpGlySerAspThrGlyThrArgLeuValThrLysPhe--GluHisAsnLeuIleAsn 609
      |||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2369 TCCGCTGGAGACTTCGGGGCGCTTGTCTCGGTTTACACGGATCGCAATTTGTTTCTGT 2310
      |||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 610 -----ArgAspGlyTyrGlnAlaGlyValAlaGluLeuArg 620
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2309 GGGCGGAGAGTGTCAATATCAAACTCAAGGTGCTTACGAAGCCATTTCGAAGGGTTCG 2250
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 621 LeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysPro----- 635
      |||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2249 CACAGCTTCATGGAATATGGGTGGAAGCTCGCTCCGTTTCCCTCGTCTCCTCTTCCCA 2190
      |||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 636 ---LeuSerHisProLeuAsnAspGlnArgAlaThrLeuGlyTyrGlnGlnGluVal 654
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2189 TTCATTCTGACGAACCGCGCGGCTACGGGATCCACGGATGGAAG----- 2139
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 655 PheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSer 674
      |||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2138 -----ATCGGTATATAATACCAGACACGTCCG-----GAGTTTGAT 2103
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 675 ArgSerIleIleGln-----AsnGlyGlyTyrAsnArgThrTyr-----Ser 688
      |||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2102 CGGTGATCTCTCCGCTCACTCAATCAATTATTCATGG---CAGACCTACTCGCACATCGT 2046
      |||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 689 LeuArgTyrArgLeuAspLysLysThrGlnAlaProGluThrTyrGlnAspLeu 708
      |||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2045 CTGGCTCATACGATCCGCTGCTGATGTCGAT-----TATCTCCATCTC 2001
      |||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 709 ProVal-----AspPheValAsnGlyLysProSerGlnGluAlaLeu----- 722
      |||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2000 CCGTATACATCGATCCCGACTCCGCCAATCCCTTCGCGCTACGACTGCATGTATAACTAC 1941
      |||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 723 -----LeuAlaGlyValAla-----ValHisLysThrValAlaAspAsnLeu 736
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1940 ACGGAGCAGTTTATCTCGCTCGGCATATATCTAGTAATATACCCGCTTCGTCCTCATG 1881
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 737 -----ValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGly 752
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1880 GAGCGTACCGTATCCAATCCT-----TTTACGGCAGCGGTTCAGTATCCAGACAGCC 1830
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 753 SerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyVal 772
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1829 GGCAACCTGCTGCAAGCCATTCTTATCTGACCGGATTCTCCGAAAGACGACACGGGTG 1770
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 773 TyrSer---PheGly----- 776
      |||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1769 TATAAAATGTTGGTCTGCACATATCTCAGTTCGTCAGCTCGATCTCGATCGGCTAAA 1710
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 777 -----AspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGly 790
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1709 ACCGTTCTTCGAAAGAGCAATACTTTG-----GCACCTGCATCTGGGTTC 1662
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 791 GlyIleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyrArgLeuArg 810
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1661 GGACTGGCTTTCCCTTAT-----GGCAATGCTCGCCATATACCCCTTTGAGTTAGCT 1611
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 811 PhePheAlaGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIle 830
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1610 TACTTTGGCGGAGGATCAACAGCGTTTCGGGCTGGAGTGTCCGATCCCTCGGCCCGGGG 1551
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 831 Ser-----AspLysGlyTyrLeuThrGly---GlyGlnValLeuAlaVal 844
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1550 AGTATGAAGATGACTCCCGCAAGACCTTCTTCGATCAGATGGGTGATATTCCTCTGGAT 1491
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 845 GlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAsp 864
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1490 CTGAATGTCGAATACAGGACAAAGCTGTCTCGAAGTTTCGGCGCAGCAGCTTTTCTCGAT 1431
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 865 IleGlyAsnAla-----Tyr 869
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 1430 GCCGCAATGTCTGCAGCATAAAGAGGATGATGAGAATCAGAGGAGGTCTCTTTGTTTC 1371
Qy 870 AspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSer 889
      |||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1370 GATCCG---TTCACAGAATAAGTATGGCTACGGTCTGGGCTTCGTCGTCGACTTC 1314
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 890 ProValGlyGlnValArgValAspValAlaThrGlyValLys----- 903
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1313 GATTATTCTTGTGCGGTGGAT-----GCCGACTGAAAGCCTACGATCCTCAGCAG 1260
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 904 -----GluGluGlyAsnProIleLysLeuHis 912
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1259 ACAGGCGCTTCAAAATGGCTATCACACGCCCAACCTTTCTTCCAATTTCTGCTTGGCAC 1200
Qy 913 PhePheIleGlyThrProPhe 919
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1199 ATTGCAGTAGGCTATCCGTTTC 1179
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 34
US-08-961-083-117
: Sequence 117, Application US/08961083
: Patent No. 6159469
: GENERAL INFORMATION:
: APPLICANT: Choi et. al.
: TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
: NUMBER OF SEQUENCES: 452
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,083
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 117:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3121 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
US-08-961-083-117

Alignment Scores:
Pred. No.: 0.00035 Length: 3121
Score: 134.00 Matches: 190
Percent Similarity: 32.04% Conservative: 124
Best Local Similarity: 19.39% Mismatches: 318
Query Match: 2.83% Indels: 348
DB: 3 Gaps: 53

US-09-914-168-2 (1-919) x US-08-961-083-117 (1-3121)
Qy 24 MetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIle-----IleAsnHis 41
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 155 TTGACCAAGCCAGATTTTATCTCCCTATAATAGTACGCTTTCTATCGGAGTCGGGAACAT 214
      :|||... ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Qy 42 ValPro-----AlaHisAspThrAlaIleAsnGlnAlaLys 53
:::|::|
Db 215 TTACCAGAGCCTCTGAAATCGAAGTTATCAATATATTTGGTTATATCAAACTAAGAAA 274
Qy 54 AlaGlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaAArgLeuAsnAlaAla 73
:::|::|
Db 275 CAGGATAATACAGAGCTTTCAAGACA-----GTTGATGGGAAATACTCTGCTCAA 325
Qy 74 GlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPheAsp----- 91
:::|::|
Db 326 AGAGATAGTCAACAACTCTACAAAACATCAGATGTAGTTTCATTGAGCTGATTTAGAA 385
Qy 92 -----GlnSerProIleSerArgIleGlyGluGlnSerProProLeuGlyLeu 107
:::|::|
Db 386 TGAACCAAGACAGAGGAGGTAGTTTACAAGGTGAAGCATCAGGGGATGATGGACTT 445
Qy 108 AspMetSerValIleGluGluThrProLeuSerLeuGluGluLeuPheAlaGlnGlu 127
:::|::|
Db 446 -----TCAGAAAATCTCTATACGACGACGAAATCTATCTCTTAATGAT 490
Qy 128 SerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnPro 147
:::|::|
Db 491 TCATTGCGCAAGTCAAGTTGACGAAAT-----CCGGATCACAAGGAGATCT--- 538
Qy 148 AsnSerGluValValProProThrLeuGluProGluLysProGlyLeuIleLysArg 167
:::|::|
Db 539 -----GTAGTTCGACCAACAGTG---CCAGAACCAAGGAAATCTCTGTCTGCT 583
Qy 168 LeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPhe 187
:::|::|
Db 583 ----- 583
Qy 188 TyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGlu 207
:::|::|
Db 584 ---ACAACGGTGCAGAGTCGGAGAGAGAGTATTGGCGACGACAAATGATCGACACAG 640
Qy 208 ProTyrAlaAsnIleLysAlaAlaLeuGluAspIle---ThrGlnGluSerAlaMetAsp 226
:::|::|
Db 641 -----TATAACTTCCTTGGAAACCAAGCGACGCAAGAACCCCGGTCTATGAG 688
Qy 227 LeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGly 246
:::|::|
Db 688 ----- 688
Qy 247 TyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIle 266
:::|::|
Db 689 -----GGTGAAGCC---GCAGTCCGT 706
Qy 267 HisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGly---Glu 285
:::|::|
Db 707 GAAGACTTA-----CCAGTCTACACT-----AGCCACTAGAACCAAGGTACACAA 734
Qy 286 GlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAsp 305
:::|::|
Db 755 GGACCCGGACATGAAGGTGAAGCTGCGGAGGAGAACACGCTTACACAGAACC 814
Qy 306 ValPhe-----HisIleGlyLysTyrGluThrLysLysAsn 317
:::|::|
Db 815 TTAGCAACGAAGACCGCAGAGCCAGGTCATGAGGGCAAGCTACAGTCCCGGAAGAG 874
Qy 318 LeuIleGlu-----AsnAlaSerAlaGluHisGlyTyr 328
:::|::|
Db 875 ACTCTAGAGTACAGCAACCGGTAGCGACAAAGGACACAAACCCGCAACAT----- 928
Qy 329 PheAspGlyArgTrpLeuAspArgSerValAspValIleLeuProAsnThrAlaAsp 348
:::|::|
Db 929 ---GAGGCG-----GAACGCGSAGTAGAAGAAGAACTTCG----- 961
Qy 349 ValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThr 368
:::|::|
Db 962 ---GCTTTAGAGTCTACTACACCAAAATAGAACCGGAAATCCAGAAATAT----- 1006

Qy 369 IleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGlu 388
:::|::|
Db 1007 -----CCTTATACAACAGAGAAATTCAGATCCACACTTCTCGAAATAATCGTCGTAAG 1060
Qy 389 LeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArg 408
:::|::|
Db 1061 ATTGAACGACAA-----GGCAAGCA----- 1081
Qy 409 AlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsn-----MetValAsn 424
:::|::|
Db 1082 -----GGACACAGCTACAATCAATATGAAGACTACATCGTAAAT 1120
Qy 425 ThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGln 444
:::|::|
Db 1121 GGTAAATGCTGTA-----GAAACT 1138
Qy 445 SerSerSerSerArgThrGluProAlaGlnValAspGlu-----SerThrLeu 460
:::|::|
Db 1139 AAAGAAAGTGTACGAGACTGAAGTAGCTCCGGTCAACGAACTCGTTAAAGTAGGAACACTT 1198
Qy 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIle 480
:::|::|
Db 1199 GTGAAAGTTAAACCTACAGTAGAAATTTACA-----AACTTAACAAAAGTT 1243
Qy 481 GluPheSerAlaSer-----AsnLeuIleGlnAspLysLeuAsnLeuVal 495
:::|::|
Db 1244 GAGAACAAAAATCTATAACTGAAGTTATAACTTAATACACACTACCTCAGCATATGTT 1303
Qy 496 AlaAlaLysAlaArgHisLeuTyrAspMetProAspArgValLeuAlaIleAsnHis 515
:::|::|
Db 1304 TCTGCAAAAACGCAA-----GTTTTCAT 1327
Qy 516 AspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAla 535
:::|::|
Db 1328 GGAGAC-----ANGCTAGTTAAA 1345
Qy 536 ArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAla 555
:::|::|
Db 1346 GAGGTGGATATAGAAAATCTCGCAAGAGCAAGTAATA----- 1384
Qy 556 LeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValProLeuTyrVal 575
:::|::|
Db 1385 -----TCAGGTTTAGATTACTACACCGGTATACAGTTAAACACACCTA 1429
Qy 576 PheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThr 595
:::|::|
Db 1430 -----ACTATAATTTGGTCAAAATAATAGGAA 1459
Qy 596 GlyThrArgLeuValThr-----LysPheGluHisAsnLeuIleAsnArgAspGly 612
:::|::|
Db 1460 AATACTGAACATCAACTCAAGATTTCCAAATTAGATAGATAAGAAAATAGAG----- 1510
Qy 613 TyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeuTyrAla 632
:::|::|
Db 1511 -----ATTAAGATATTGATTTCAGTACAGTATTAACGGT 1543
Qy 633 ThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGln 652
:::|::|
Db 1544 AAAGAA-----AATGATCGTTATCGTAGATATTTAAGTCTTAAGT--- 1582
Qy 653 GluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGlu 672
:::|::|
Db 1583 -----GAAGCCCGGCTGATACGGCTAAA----- 1606
Qy 673 IleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThrTyrSerLeuArgTyrArg 692
:::|::|
Db 1607 -----TACTTTCTAAAAGTGAAA 1624
Qy 693 LeuAspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPhe 712
:::|::|
Db 1625 TCAGATCGCTTCAAAAGAAATGTAC-----CTACCTGTAAAATCT 1663
Qy 713 ValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrVal 732

```
Db 1664 ATTACA---GAAATAGCATGGTGAACAGTATATAAAGTACGGTAGCGTGTGATCAACTGTGC 1720
Qy 733 AlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGly 752
Db 1721 GAAGAAGGT-----ACAGCGGTTACAAAGATGATTACACATTTACTGTGTAGCT 1768
Qy 753 SerSer-----GlyLeuValSerAspAlaAsnMetAlaIleAla----- 765
Db 1769 AAATCTAAAGCAGACGACCAAGGAGTTTACACATCCTTTAAACAGCTGGTAACAGCCATG 1828
Qy 766 ArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsn-----Ala 779
Db 1829 CAAAGCAATCTGTGTGTCTATACATTTGCTTACAGATATGACCGCAGATGAGTGAAGC 1888
Qy 780 TyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrSer 799
Db 1889 TTAGCGCATTAAGCAGACAAAGTTATCTCACAGGTCCATTACAGCGAGCTTCATCGGTCT 1948
Qy 800 AspAsnPhe---AsnHisValProTyrArgLeuArg-----PhePheAla 813
Db 1949 GATGGAACAAATCGTATGCCATTTATGATTTGAAGAAACCATTTATTGATACATTAAT 2008
Qy 814 GlyGly-----AspGlnSerIleArgGlyTyrAlaHisAspSer----- 826
Db 2009 GGTGCTACAGTTAGAGATTTGGATATTAAAACTGTTTCTGCTGATAGTAAGAAAAATGTC 2068
Qy 827 -----LeuSerProIleSerAspLysGlyTyr 835
Db 2069 GCACCGCTGGCAGGAGCGCAATAGCGCGAATATTAAATATGTTGCGTAGAAGGAAAA 2128
Qy 836 LeuThrGlyGlyGlnValLeu-----AlaValGlyThrAlaGluTyr 849
Db 2129 ATCTCAGGTGCGAAATCTGTTGCGGATTTAGTAGCGAGCGCAACAAATACAGTATAGAA 2188
Qy 850 AsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr 869
Db 2189 AACAGCTCGTTTACAGGGAAACTT-----ATCGCAAAATCACCAG 2227
Qy 870 AspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGly-----ValArgTyr 887
Db 2228 GACAGTAAT---AAAATGATACT-----GGAGGAATAGTAGTAATATACAGGAAT 2278
Qy 888 AlaSerProValGlyGlnValArgValAspValAlaThrGlyValLysGluGlyAsn 907
Db 2279 AGTTTCGAGAGTTAATAAAGTTAGGTAGATGCCTTAATCTCTACTAATGCACGCAATAAT 2338
```

RESULT 35

US-08-770-379-20

```
; Sequence 20, Application US/08770379
; Patent No. 5849564
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
; TUMOR VIRUS, DNA ENCODING SAME AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/770,379
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 52342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32207 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-770-379-20
```

```
Alignment Scores:
Pred. No.: 0.0302 Length: 32207
Score: 132.50 Matches: 167
Percent Similarity: 33.29% Conservative: 118
Best Local Similarity: 19.51% Mismatches: 300
Query Match: 2.80% Indels: 271
DB: 2 Gaps: 39
```

```
US-09-914-168-2 (1-919) x US-08-770-379-20 (1-32207)
Qy 6 LeuPheAlaAsnArgSerPheMetProValAlaLeuAlaAlaTyrLeuProLeuMetThr 25
Db 1514 CTGTATCTCCGCAATCATTTATCAGCGTGACCCCTA----- 1549
Qy 26 SerGlnAlaLeuAlaGlnGlnAsnProAlaAsnIleLeuAsnHisValProAlaHis 45
Db 1550 -----CACAGTACAAACCTTGGAAACCTTAGTTACCAAGCTACCA----- 1588
Qy 46 AspThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGluGln 65
Db 1589 -----AAACTTGTGTAGGCGCTTCCCGGGGCGCCGACCTCTCTGACAAGCCG----- 1636
Qy 66 IleGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAsp 85
Db 1637 ---TCCGTGAGACAGTCCCTTTCCACCCCTGTGCACAAACCTGTCGCGAGATGCCCTGGAC 1693
Qy 86 ValValAsnPheAspAspGlnSerProIleSerArgIleGlyGluGlnSerProProLeu 105
Db 1694 GCCCTG-----GAAAAAAGGATCCG--- 1714
Qy 106 GlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluGluPheAla 125
Db 1715 -----GCCCTTCTTGTGAGGGGACCACGTTGGCGCTGGAGACACTCTCTAGGA 1762
Qy 126 GlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGlu 145
Db 1763 TACGGGTGCGTGCAG-----GACTACAAGGAGACGGTACAG----- 1798
Qy 146 GlnProAsnSerGluValValProProThrLeuGluProGluLysProGlyLeuIle 165
Db 1799 -----ATAATATCCAGCCTTGTGGCGCATCCAAAA-----TTAGTCT 1834
Qy 166 LysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLysAla 185
Db 1835 AGGACACCGGCGCGACAGTGGCCACTGCCGTGACAAGCTAACTGACCTCAATCA 1894
Qy 186 -----LysPheTyr----- 188
Db 1895 ACTGTGCCACGACGCCCATCGACAGCGCTACGAAACGAAATATACAGATTGATCA 1954
Qy 189 -----GlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLys 205
Db 1955 AGGACCTCAAGAGGCTCAAAAACACGAGACC-----AATCGGCC 1996
```

Qy	206	ThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAlaMet	225
Db	1997	ATGGAGGAATGGAGCAGAGAAGTACTTGGCTCTTGCAATGGCTCTCCGGAACGTGTCCGC	2056
Qy	226	AspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArg	243
Db	2057	ACCTCTC-----CTGCACACAGGCTCCACCGCGAAGCTTAGAGATTT	2098
Qy	244	AlaValGlyTyrTyrAspIleAspLeuSerIle-----IleArgAsnSer	258
Db	2099	GCAGAGAACACTTCANAATACTACTCCCGGTACCCGGGACGCCCGCTCAACAGCTCT	2158
Qy	259	IleGlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArg	278
Db	2159	CCAAAGCCGATGGAATACAGCGCACCCCTCCCGGACCCCAAGGATATAGC---AGA	2215
Qy	279	AlaValGluValArgGlyGluGlyAlaAspAspLys-----	290
Db	2216	GCTACATCCATCCACGGGGAACAGGCGTGGAGAAAGATACACAGAGCGCTTCAAGGATTTC	2275
Qy	291	-----AlaPheThrThrValAlaAsp-----	297
Db	2276	AACITTCGCCCTTCGGCCCGCTGACTGGATGCCCTGGCAGCGGAGTACCAACGCCGT	2335
Qy	298	-----GluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLys	315
Db	2336	GGTTACGCCCTTCGGCGCGCGTGGGTCCACGGCGCTC---TCAGGCTTCCTGGAGACGATC	2392
Qy	316	LysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAsp	335
Db	2393	CTAGGGAGCGTGAACGACATC-----TACATGGAT-----	2422
Qy	336	ArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThr	355
Db	2423	--AAGCTCCGCTCCTTCCTGCCCGCAGCGCAGCCTTTTCAGCGCGCCCTTCGACTGG	2479
Qy	356	GlyThrGlnTyrArgPheAspGluValValPhePhe-----ThrIleAspProLysThr	373
Db	2480	CTAAGCGCGTATCAG---GACCAGGTACAGCTTTTCTTCGCACCACCATAGGG-----	2527
Qy	374	AsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeu	393
Db	2528	-----CTGCGCTGGTGGCGAGCGCTGGCGGACCAAGATC	2560
Qy	394	LeuThrValAsnMetGlyGluAlaTyrAsnLeu-----GlnAlaVal	407
Db	2561	AGCGTGCGAGCACTGAGGCTTACCCACGCGCTCCAGTCCGGCGATTTCACGAGCCACG	2620
Qy	408	ArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetVal---AsnThrGlu	426
Db	2621	GTGGGCACGCCCTGGAGCTCCCTGTCACACAGATAGCGCGCATCGCTCCCAACATGAAG	2680
Qy	427	IleValPheProGluArgGlu---GlnIleGlnAsnAspGlnValSerPheGluGlnSer	445
Db	2681	TCCGTGTTCAACGACCACGAGCTTACGTTGCGCATCAGAGTTCGGGATTATGTGGAGGCC	2740
Qy	446	SerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGlu	465
Db	2741	CAACGCGCGACGCACACGCCACACGCTCCCGCTCCCAAGATACAGCGACCAAAAGACT	2800
Qy	466	ThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSer	485
Db	2801	CTGATCTCACATCCGAGCAATCGTCGCGGACGGCATCC---GCCTTTCTTTAAGACG	2857
Qy	486	AsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMet	505
Db	2858	TCCCTACTGCAGCAAGAGCCAACTTCTGGCGCTACAG-----	2896
Qy	506	ProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGly	525
Db	2896	-----	2896
Qy	526	ArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsn	545

D	b	2897	- - - - - : : : : : - - - - - CGGGCCGAGTC	2900
Q	y	546	GluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspVal	565
			: : : : : : : :	
D	b	2909	GAGTGCCTCAGACCGACATGCCGCCGAGAGGCCAGAGAAGCATCCGCGAGAA	2968
Q	y	566	TyrGlnSerLys- - - - - LysValProLeuTyrVal	575
			: : : : : : : : : : : : : : : : : :	
D	b	2969	ACCCAGCGCAAAATGGCACACGCCATCACTCAGCTCTTACAGCAGGCACCCAGTGCGCATC	3028
Q	y	576	PheValAlaSerAspLysProArgaspGlyGlnIleGlyLeuGlyTyrPglySerAspThr	595
D	b	3029	- - - - - TCGGGGCGCCGTACTCTTACAGACCGCGTGGGCTC- - - - -	3067
			: : : : : : : :	
Q	y	596	GlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAla	615
D	b	3068	- - - - - CTCGAGGGCATCATACGACAAGGTCCTGGAGCGGAANTCTACGAGACG	3118
			: : : : : : : : : : : : :	
Q	y	616	GlyAlaGlu- - - - - LeuArgLeuSerGluAspLysLysGlyValLysLeuTyrAla- - - - -	632
			: : : : : : : : : : : : :	
D	b	3119	GGTCTCAGAGGACTGTCTCGCTCGAGCAGACCATTCAAGTCATCACCCTATACGCTCCC	3178
Q	y	633	- - - - - ThrLysProLeuSerHisProLeuAsnAspGlnLeu- - - - - ArgAlaThr	647
			: : : : :	
D	b	3179	GTACAGAGAAAGCAATGCACGTCTGCTCGAGAGGTGAAAACACCGAGCAAAC	3238
Q	y	648	LeuGlyTyrGlnGlnValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArg	667
D	b	3239	ACTGAGACCGCTCTCGAGCTAGAGCGCGGTACGCACGCGCAGCAC- - - - - GCTAGA	3292
			: : : : : : : : : :	
Q	y	668	ThrLeuGluHisGluIleSerArg- - - - - SerIleIleGlnAsnGlyGlyTrpAsnArg	685
			: : : : : : : : : :	
D	b	3293	CTCTCTGAGCGAGCGGTGCATGAGCTGTCCACCGTTTAGGGGGGGAAGCCGCG	3352
Q	y	686	ThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrp	705
D	b	3353	GTGCAATCCTCGCGGCAAAAAATCCAACCCCTGMAATCCCTGGTA- - - - -	3397
			: : : : :	
Q	y	706	GlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGly	725
D	b	3398	- - - - - CAGAAAGCGGAGCAGCGCCGC	3418
			: : : : :	
Q	y	726	- - - - - ValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyr	743
D	b	3419	CTCCTGTTGGCCACCATAGACCGTGGCC- - - - -	3448
			: : :	
Q	y	744	ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla	763
D	b	3449	- - - - - GGCCAGGCCGAGGAGACC	3466
Q	y	764	IleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAla	779
			: : : : : : : :	
D	b	3467	ATATCACCATTCCACTCCAGGACTGTACCAACAGGGACAGGAGGCC	3514
			: : : : :	
		RESULT 36		
		US-08-757-669A-20		
		; Sequence 20, Application US/08757669A		
		; Patent No. 6183751		
		; GENERAL INFORMATION:		
		; APPLICANT: Chang, Yuan		
		; APPLICANT: Bohenzky, Roy A.		
		; APPLICANT: Russo, James J.		
		; APPLICANT: Edelman, Isidore S.		
		; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS		
		; TITLE OF INVENTION: SEQUENCES AND USES THEREOF		
		; NUMBER OF SEQUENCES: 20		
		; CORRESPONDENCE ADDRESS:		
		; ADDRESSEE: Cooper & Dunham LLP		
		; STREET: 1185 Avenue of the Americas		
		; CITY: New York		
		; STATE: New York		

RESULTS 36

```

RESOLUTION 20
US-08-757-669A-20
; Sequence 20, Application US/08757669A
; Patent No. 6183751
; GENERAL INVENTION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York

```

COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/757,669A
 FILING DATE:
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 45185-F
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391-0525
 INFORMATION FOR SEQ ID NO: 20:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 32207 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-08-757-669A-20

Alignment Scores:
 Pred. No.: 0.0302 Length: 32207
 Score: 132.50 Matches: 167
 Percent Similarity: 33.29% Conservative: 118
 Best Local Similarity: 19.51% Mismatches: 300
 Query Match: 2.80% Indels: 271
 DB: 4 Gaps: 39

US-09-914-168-2 (1-919) x US-08-757-669A-20 (1-32207)

```

Qy 186 -----LysPheTyr----- 188
Db 1895 ACTCTGCCACGACCGCATCGAGACGCTAGAAACGAACTATACAGATTGATCAA 1954
Qy 189 -----GlnSerSerGlnSerGlyCluThrSerAlaIleGlySerSerHisGlnLys 205
Db 1955 AGGACCTCAAAGAGGCTCAAAACACAGAGACC-----AATCGGCCC 1996
Qy 206 ThrGluProTyrAlaAsnIleLysAlaAlaGluAspIleThrGlnGluSerAlaMet 225
Db 1997 ATGGAGGAATGAACGAAAGTACTGGCTCTTACAAATGCTCTCCGAGACGTGTCGCC 2056
Qy 226 AspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuAlaAlaArg----- 243
Db 2057 ACCCTC-----CTCAACAGGCTCCACCGCGAGGCTTAGAGAGTTT 2098
Qy 244 AlaValGlyTyrTyrAspIleAspLeuSerIle-----IleArgAsnSer 258
Db 2099 GCAGAGACGACTTCAAAATATCTACTCCCGTACCCGCGGACGCCCGCTCCAAGCGTCT 2158
Qy 259 IleGlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArg 278
Db 2159 CCAACGCGGATGAATACAGCGCCAGCCCTCCCGGACCCAAAGGATATAGAC---AGA 2215
Qy 279 AlaValGluValArgGlyGluGlyAlaAspLys----- 290
Db 2216 GCTATCTCCACGCGGAACAGCGCTGGAAGAAGATACAGCAGCGCTTCAGGATTTTC 2275
Qy 291 -----AlaPheThrThrValAlaAsp----- 297
Db 2276 AACTTCGCGCTCTGCGCGCGCTGACTGGGATCCCTGGCAGCGGAGTACCAACGCGGT 2335
Qy 298 -----GluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGlnLys 315
Db 2336 GGTTCGCGCTCTGCGCGCGCTGGTCCAGCGCTC---TCAGGGTCTCTGAGACGATC 2392
Qy 316 LysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAsp 335
Db 2393 CTAGGGACGCTGAACGACATC-----TACATGGAT----- 2422
Qy 336 ArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThr 355
Db 2423 ---NAGCTCCGCTCTTCTGCGCGCGCGAGCTTTTCAGGCGCCCGCTTCGACGTGG 2479
Qy 356 GlyThrGlnTyrArgPheAspGluValValPhePhe-----ThrIleAspProLysThr 373
Db 2480 CTAACGCGGTATCAG---GACCAAGTCAGCTTTTCTTGGCGCACCATAGG----- 2527
Qy 374 AsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGluGlnLeu 393
Db 2528 -----CTGCGCTGCTGCGAGCGCTGCGCGGCTGCGCGACAGATC 2560
Qy 394 LeuThrValAsnMetGlyGluAlaTyrAsnLeu-----GlnAlaVal 407
Db 2561 AGCTGGCAGGCACTGAGGCTTAGCCAGCGCTCCAGTCCGCGGATTTGACAGCGGCCACG 2620
Qy 408 ArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetVal---AsnThrGlu 426
Db 2621 GTGGCAGCGCCCTGGAGCTCCCTGCCACAGTACGCGCGCATCGCTCCCAACATGAAG 2680
Qy 427 IleValPheProGluArgGlu---GlnIleGlnAsnAspGlnValSerPheGluGlnSer 445
Db 2681 TCCGTGTTCACAGCAGCGGACTTCAGGTGCGATCAGAGGTGCGGATATATGTGAGGCC 2740
Qy 446 SerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGlu 465
Db 2741 CAACGAGCGGACGACACCGCTCCACGTCCCAAGATACAGGACCCCAAGACT 2800
Qy 466 ThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSer 485
Db 2801 CTGATTCACATCCGGAGCGCAATCGTCGCGGACGAGTACCC---GCCTTTCTTAAGACG 2857
Qy 486 AsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMet 505

```

```
Db 2858 TCCCTACTGCACAGAGCCCAAACTTGGCGCTACAG----- 2896
Qy 506 ProAspArgValLeuAlaIleAsnHisAspGlyValAsnArgSerIleLeuGly 525
Db 2896 ----- 2896
Qy 526 ArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsn 545
Db 2897 -----CGGGCGGACTTC 2908
Qy 546 GluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspVal 565
Db 2909 GAGTCGCTCGAGGCGACATCGGCCCGCAGAGCCAGAGAAAGCATCCGCGCAGGAA 2968
Qy 566 TyrGlnSerLys-----LysValProLysVal 575
Db 2969 ACCCAGCGCAAAATGGACACGCCATCCTCAGCTCTTACAGCAGGACCCAGTCGCATC 3028
Qy 576 PheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAspThr 595
Db 3029 -----TCGGGGCGCCCTACTCCTTACAGGACCGCGTGGCTTC----- 3067
Qy 596 GlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAla 615
Db 3068 -----CTCAGGCGCATCATATACGACAAAGTCTCTGGAGCGCGAATCCTACGAGCG 3118
Qy 616 GlyAlaGlu-----LeuArgLeuSerGluAspLysLysGlyValLysLeuTyrAla--- 632
Db 3119 GGTCTCAGGAGGACGTCTCTGGCTCGAGCAGACCATCAAGTCCATCACCGTATACGCTCCC 3178
Qy 633 -----ThrLysProLeuSerHisProLeuAsnAspGlnLeu-----ArgAlaThr 647
Db 3179 GTAGAGGAGAGCAAAAGTGCAGCTGCTGTGACGAGGTGAAAGACGACGAGCAAAAC 3238
Qy 648 LeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArg 667
Db 3239 ACTGAGACCGCTCTCGAGCTAGAGCGCGCTACGACGCGCAGCAGAC-----GCTAGA 3292
Qy 668 ThrLeuGluHisGluIleSerArg-----SerIleIleGlnAsnGlyGlyTrpAsnArg 695
Db 3293 CTCTCGAGGCGAGGTCGATGAGCTGCTACCGTTCACCGTGAAGGGGGGAAAGCGCGC 3352
Qy 686 ThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrp 705
Db 3353 GTGGAATCCTGGCGGAGAGAAATCCAAACCTCAATCCCTGTA----- 3397
Qy 706 GluAspLeuProValaspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGly 725
Db 3398 -----ValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyr 743
Qy 726 -----ValAlaValHisLysThrValAlaAspAsnLeuValSerAspAlaAsnMetAla 763
Db 3419 CTCCTGTGGCCACCATAGACAGTGGCC----- 3448
Qy 744 ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
Db 3449 -----GCCCAGGCCCGCAGGAGACC 3466
Qy 764 IleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAla 779
Db 3467 ATATACCATCCACATCCAGGAGCTGTACCAACAGGAGGAGGCC 3514
```

RESULT 37

```
US-09-230-371A-20
; Sequence 20, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
```

```
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
; FILE OF INVENTION: US97/230, 371A
; CURRENT APPLICATION NUMBER: US/09/230, 371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 32207
; TYPE: DNA
; ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-20
```

Alignment Scores:

```
Pred. No.: 0.0302 Length: 32207
Score: 132.50 Matches: 167
Percent Similarity: 33.29% Conservative: 118
Best Local Similarity: 19.51% Mismatches: 300
Query Match: 2.80% Indels: 271
DB: 4 Gaps: 39
```

US-09-914-168-2 (1-919) x US-09-230-371A-20 (1-32207)

```
Qy 6 LeuPheAlaAsnArgSerPheMetProValAlaLeuAlaAlaTyrLeuProLeuMetThr 25
Db 1514 CTGGTATCTCCGAATCATTTATCACCCTGACCCCTA----- 1549
Qy 26 SerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleAsnHisValProAlaHis 45
Db 1550 -----CACAGTACAAACCTTGGAAACCTATTACCAACCTACCA----- 1588
Qy 46 AspThrAlaIleAsnGlnAlaLysAlaGlyAsnProValLeuLeuThrProGluGln 65
Db 1589 -----AAACTTGGTGAGCGGTTTCCCGGGGGCCGACCTCTCACAAGCCG----- 1636
Qy 66 IleGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAsp 85
Db 1637 ---TCGCTGAGACAGTCCCTTTCCACCTGTGCACACCCCTCGCGAGATGCCCTGGAC 1693
Qy 86 ValValAsnPheAspAspGlnSerProIleSerArgIleGlyGluGlnSerProLeu 105
Db 1694 GCCTTG-----GAAAAAAGGATCCG----- 1714
Qy 106 GlyLeuAspMetSerValIleGluThrThrProLeuSerLeuGluLeuPheAla 125
Db 1715 -----GCCCTTCTGTCAGGGGACCACTGGCGCTGGAGACACTCCTAGGA 1762
Qy 126 GlnGluSerThrGluMetClyIleAsnProAsnAspTyrIleProGluTyrGlnGly 145
Db 1763 TACGGGTCCGTGCAG-----GACTACAAGGAGACGGTACAG----- 1798
Qy 146 GlnProAsnSerGluValValProThrLeuGluProGluLysProGlyLeu 165
Db 1799 -----ATATATCCACCCCTTGGCGCATCCCAAAA-----TTAGTC 1834
Qy 166 LysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAla 185
Db 1835 AGGGACCAAGCGCGGACCAAGTGGGCCACTGCCGTGACAAGGCTAACTGACCTCAATCA 1894
Qy 186 -----LysPheTyr----- 188
Db 1895 ACTCTGGCCACACCGGCATCGACAGGCTACGAAACGGAATATACAGATTGATCCAA 1954
Qy 189 -----GlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLys 205
Db 1955 AGGACCTCAAGAGGCTCAAAACACGAGACC-----AATCGGGCC 1996
Qy 206 ThrGluProTyrAlaAsnIleLysAlaLeuGluAspIleThrGlnGluSerAlaMet 225
Db 1997 ATGGAGGAATGAAAGCAGCAAGAGTACTGTGCTTTCACAATGCTCTCCGGAACGTTGCC 2056
```

Qy 226 AspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArg----- 243
Db 2057 ACCCTC-----CTGCAACAGGCTCCACCAGCGCTAGAGAGTTT 2098
Qy 244 AlaValGlyTyrTyrAspIleAspLeuSerIle-----IleArgAsnSer 258
Db 2099 GCAGAGAACACTTCAAATACTACTCCCTCCCGCGGACGCCCGCTCCCAAGCGTCT 2158
Qy 259 IleGlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArg 278
Db 2159 CCAACGGCGATGGAATACAGCGCCAGCCCTCCCGGACCCCAAGGATATAGC---AGA 2215
Qy 279 AlaValGluValArgGlyGluGlyAlaAspAspLys----- 290
Db 2216 GCTACATCCATCCACGGGGAACGGCGTGGAAGAAGATACAGCAGCGCTTCAAGGATTTC 2275
Qy 291 -----AlaPheThrValAlaAsp----- 297
Db 2276 AACTTCGCGCGTCTGCGGCCCGCTGACTGGGATGCCCTGGCAGCGGATACCAACGCCGT 2335
Qy 298 -----GluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLys 315
Db 2336 GGTTCGCCCTTCCGGCGCGTGGTCCAGCGCTC---TCAGGGTTCCTGGAGACGATC 2392
Qy 316 LysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAsp 335
Db 2393 CTAGGGAGCGTGAACGACATC-----TACATGGAT----- 2422
Qy 336 ArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThr 355
Db 2423 ---AAGCTCCCGCTCTTCTGCCGACCGCGAGCGCTTTTCAGCGCGCCCTTCGACTGG 2479
Qy 356 GlyThrGlnTyrArgPheAspGluValValPhePhe-----ThrIleAspProLysThr 373
Db 2480 CTAACGCGGTATCAG---GACCAAGTCAAGCTTTTCTTGGCCACCATAGG----- 2527
Qy 374 AsnGlnLeuThrAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeu 393
Db 2528 -----CTGCGCTGGTGGCAGCGCTGGCGCGCAAGAATC 2560
Qy 394 LeuThrValAsnMetGlyAlaTyrAsnLeu-----GlnAlaVal 407
Db 2561 AGCGTGACAGGACTGAGGCTTAGCCACGCGCTCCAGTCCGCGGATTTCCACGAGGCCACG 2620
Qy 408 ArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetVal---AsnThrGlu 426
Db 2621 GTGGCAGCGCCCTGGAGCTCCCTGCCACAGAGTACGGCGCATCGCCTCCAACATGAAG 2680
Qy 427 IleValPheProGluArgGlu---GlnIleGlnAsnAspGlnValSerPheGluGlnSer 445
Db 2681 TCCGTGTTCAACAGCACCGGACTTCAGGTGCGATCAGAGGTGCGGGATTATGTGGAGGCC 2740
Qy 446 SerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGlu 465
Db 2741 CAACGAGCGGACGACACACCCACACGCTCCACGTCCTCAAGATACAGCCACCAAGACT 2800
Qy 466 ThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSer 485
Db 2801 CTGATTCCACATCCGGACGCAATCGTCGGACGCGACTACCC---GCCTTTCTTAAGACG 2857
Qy 486 AsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMet 505
Db 2858 TCCCTACTGCAGAGGCGCAAACTTCTGGCGGTACAG----- 2896
Qy 506 ProAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGly 525
Db 2896 ----- 2896
Qy 526 ArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsn 545
Db 2897 -----CGGCGGCGACTTC 2908
Qy 546 GluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspVal 565

Db 2909 GAGTCGCTCGAGAGCGACATCGCGCGCCGACAGAGGCCAGAGAAAGCATCGCGCAGGAA 2968
Qy 566 TyrGlnSerLys-----LysValProLeuTyrVal 575
Db 2969 ACCGAGCGCAAAATGGACACGCCCATCTCAGCTCTTTACAGCAGCCAGCATGCGCATC 3028
Qy 576 PheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrPheLysAspThr 595
Db 3029 -----TCGGGCGCGCGCTATCTCTACAGGACCGGTGGGCTTC----- 3067
Qy 596 GlyThrArgLeuValThrLysPheGluHisAsnLeuLeuAsnArgAspGlyTyrGlnAla 615
Db 3068 -----CTCAGGGCATCATATACAGAGGTCTCTGGAGCGCAATCTCTAGCAGACG 3118
Qy 616 IleAlaGlu-----LeuArgLeuSerGluAspLysLysGlyValLysLeuTyrAla--- 632
Db 3119 GGTCTCAGGAGACTGTCTCTGCTCAGCAGACCATCACTCCATCACCCTATACCTCC 3178
Qy 633 -----ThrLysProLeuSerHisProLeuAsnAspGlnLeu-----ArgAlaThr 647
Db 3179 GTAGAGGAGAAACAAAGAAATGCACGTCTGCTGGACGAGGTGAAAAGCAGCAGCAAC 3238
Qy 648 LeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArg 667
Db 3239 ACTGAGACCGCTCTCGAGCTAGAGCGCGCTAGCGACGCGCGAC-----GCTAGA 3292
Qy 668 ThrLeuGluHisGluIleSerArg-----SerIleLeuGlnAsnGlyTyrAsnArg 685
Db 3293 CTCCTGCAGCAGCGGTGCTGATGAGTGTACCGTTCGCCGTTAAGGGGGGAAGCCCGG 3352
Qy 686 ThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrp 705
Db 3353 GTGGAATCTCGCGCAGAAAATCCAAACCTGAAATCCCTGGTA----- 3397
Qy 706 GlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGly 725
Db 3398 -----CAGGAAGCGGAGCAGCGCGC 3418
Qy 726 -----ValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyr 743
Db 3419 CTCCTGTGGCCACCATAGACACGGTGCC----- 3448
Qy 744 ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
Db 3449 -----GGCAGGCCAGGAGGCC 3466
Qy 764 IleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAla 779
Db 3467 ATATCACCATCCACTCCAGGACTGTACCAACAGGAGGAGGCC 3514

RESULT 38

US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: TUBERCULOSIS
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ. ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:

```

Pred. No.:      211      Length:      4411529
Score:          131.50   Matches:      165
Percent Similarity: 33.21%  Conservative: 102
Best Local Similarity: 20.52% Mismatches: 276
Query Match:      2.78%   Indels:    262
DB:              4       Gaps:      36

US-09-914-168-2 (1-919) x US-09-103-840A-1 (1-4411529)

QY 15 ValAlaLeuAlaAlaTyrLeuProLeuMetThrSerGlnAlaLeuAlaGlnAlaAsn 34
|||||
DB 3496875 GTCCCTTGGCGGCGCATACCGCGCGACCGACCGATCGCGCGCTACAGTTGCTGG 3496934
QY 35 ProAlaAsnIleAlaAsnHisValProAlaHisAspThrAlaIleAsnGlnAlaLysAla 54
|||||
DB 3496935 CCG-----GTCCAGAGGTGACATCCCGTGTGTCGCCGCAAGAGCA 3496976
QY 55 GlyAsnProProValLeuLeu----- 61
DB 3496977 GCCCGTGGCGGGTCAATGCTCGCGCAGTCAGTCGCCGTAGACCGCAATGAAGCCATTGTG 3497036
QY 62 -----ThrProGluGlnIleGlnAlaArg----- 69
DB 3497037 GCGCAGGCTGCTTGGACACAGTGCACCGAGGGA-ATATCTCGCGAGCTACCAATTGG 3497095
QY 70 -----LeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspVal 86
|||||
DB 3497096 GAGTGGGCGCTAGGTTTCAGTGGCGGGTTCGCCGCCGCAAGAGCGCCATCAGACCC 3497155
QY 87 ValAsnPheAspGlnSerProIleSerArgIle-GlyGluGlnSerProProLeuG1 106
|||||
DB 3497156 CAGTGGCCCGATCCCGGTCGCTGTTCGCCGGCGCGGTGTCATCAGCCGCTCC----- 3497210
QY 106 yLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluGluPheAlaG1 126
|||||
DB 3497211 ----GAGCTTTACCGCTGACGACGCGCGCGGATC----- 3497243
QY 126 nGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluG1 146
|||||
DB 3497244 -----CTGGCCTAGGACCGAGACAGACGACGCGGTGGCGCGCGCGGGA 3497296
QY 146 nProAsnSerGluValValValPro-----ProThrLe 157
|||||
DB 3497297 GGCGCCGACGATCGTCTTGTTCACGCGACGCAATGGGGCTGGCTGCTGCCCATCAC 3497356
QY 157 uGluProGluLys----- 161
DB 3497357 CGAACCGCTGGAGATCGCCAAAGCCGCGACCGCGTCCGTGCGGAGTGTTCGCGCGCGG 3497416
QY 162 -----ProGlyLeuIleLysArgLeu---TyrAlaArgLeuPheAsnAspGlyValAs 178
|||||
DB 3497417 CGGTACCCCGCAGATGCTGCTCGAGTGGTGGCAGCGATCAATGCGACCGCTGCC 3497476
QY 178 nLysValProArgLeuLysAlaLysPheTyrGlnSer-SerGlnSerGlyGluThrSerA 198
|||||
DB 3497477 ACCGACGCCACCG-----CGCGAACTGTCCAGGTCTGTGAGTGGCGCGGAGAGCT 3497527
QY 198 lalle-----GlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaL 216
|||||
DB 3497528 ACTGGACACACGGGTGTGACCATCGACGACTGTTCGCTCGCG----- 3497571
QY 216 euGluAspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgG 236
|||||
DB 3497572 -----CCCGTACGCTCGGAGGGTGAATT 3497596
QY 236 lnThrAlaLeuValAlaAlaArg-AlaValGlyTyrTyrAspIleAspSerIleIle 255
|||||
DB 3497597 GCCCGCCGCGCTGCTCTCGCGCTGCCGAGGT-----TCGTTAAGCGCTTC 3497644
QY 256 Arg-AsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGluProValIlyl 275
|||||

```

```

DB 3497645 CGGTGAACCTCGTCGGGCGACGCCCGCGCGCTTGTTCGGTGCACCTCGATGCACAACTCCTCCT 3497704
QY 275 eAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrVa 295
|||||
DB 3497705 CGACTTT-----GACCGGACGGTCAAGCTGCTGCCCTTCCGATCGCGACCGCGT 3497755
QY 295 lAlaAspGluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrly 315
|||||
DB 3497756 TGTGTACCGCTTCCCG-----AACCCAC-----CG 3497779
QY 315 sLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgThrLeuAs 335
|||||
DB 3497780 CCTCGCCTGATCGCGACGCGCTGCTCGACCAC-----CGACAATGGACCCA 3497827
QY 335 pArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspTh 355
DB 3497828 CGAA-----TTCAACGC 3497839
QY 355 rGlyThrGlnTyrArgPheAspGluValValPhePheThrIleAspProIlyThrAsnG1 375
|||||
DB 3497840 TGTGCGCNAACCCGACGTCGCGCAATTTGGCTACGCGCGCATC----- 3497882
QY 375 nLeuThrThrAspProAspLysLeuProValLysArgGluLeuGlu-----GlnLeuLe 394
DB 3497883 -----GATTCTCCTCCGAGCGCGAGTGATACCTCGCATGCTCCGTCCA 3497926
QY 394 uThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLe 414
DB 3497927 GGTCAATAATGGTGGTCCGGATTTC----- 3497951
QY 414 ulleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluG1 434
DB 3497952 -----CTGTATAACGCTCTG-----CAGATCTCTACCACTCCCGAGAGTC 3497992
QY 434 nIleGlnAsnAspGlnValSerPheGlu-----GlnSerSerSerSerAr 449
DB 3497993 GTTCTGTCACCTTCAGGATTACGTTTCGTGCGGGACAGCACCTTCGAAGCCAGGCAATCG 3498052
QY 449 gThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLe 469
DB 3498053 CGAAGACCG-----CTGGATGACATGTCATGGAGGTTCAGCGCGCATACCATCCCGGT 3498106
QY 469 uThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleG1 489
DB 3498107 CGGT-----CAGTACGTCTGCGCATCCGACGCTGCGAAG 3498145
QY 489 nAspLysLeu-----AsnLeuValAlaAlaLysAlaArgHisLeuTyr----- 503
DB 3498146 TGGCCAAATTGCCAAGCCAGCGCGCTGTCGGGCAACGGCGCATCATCTCGAGTTGTT 3498205
QY 504 -----AspMetProAspAspArgValLeuAlaIleAsnHisAsp----- 516
DB 3498206 CGTCGGTGAAAGCCCTCGACCGCTTGACTCAGCAACAACAGCACACCGCTACGGTAC 3498265
QY 517 -AspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaAr 536
DB 3498266 CTCTGGCGCGGACGCAACAGCAGCGCGGACCTCGCTCGCGCAGTTCGTCAGGCGCTT 3498325
QY 536 gAlaIleLeuPro-----AspGluSerGluAsnGluVal---IleAspLeuProGluArgTh 554
DB 3498326 CCAATCGACCGCTCGACCGCTCGCGGAATGCGGTTCAGCAAGACCTCCCGGACCGC 3498385
QY 554 rAlaLeuAlaAsnArgLysThrPro-----AlaAspValTyrG1 567
DB 3498386 CGCCCGCCCGCAATGTTTCGCGCCCAACAATGGAAGCCACAGCGCTGCCGACTGTTTCAA 3498445
QY 567 nSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgaspGlyGlnI1 587
DB 3498446 TCACCAGCAGCTCCCC----- 3498461
QY 587 eGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLe 607
DB 3498462 -----CAGTCAGCGCGGACGATGCTCCT 3498484

```


[illegible]

US-09-134-001C-224

Db	6733	ATTGTAAACCAAGGCTTAATGAATAATTATGAAAAAAGCTGCGGAAACAAAGCACTTTA	6792
Qy	227	LeuAsn-----GlySerIlePro	232
Db	6793	ATAAATCAAAATCAAGATCGACACTAGAAAGAAACAAATAGCACTTGGCAAAATAGAA	6852
Qy	233	ArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeu	252
Db	6853	GMACTAAAGAATCAAGCGTTAAATCAAGTA---TCACAGGCACACATCAAAATGAATGATGTG	6909
Qy	253	SerIleIleArgAsnSer-----IleGlyGluVal-----AspValIleIle	266
Db	6910	AAATTCGCGAAATTAATGGAATTGCTAAATTTCTGAGGTCCATCTTGAGACTATAATT	6959
Qy	267	-----HisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGlu	281
Db	6970	AAACGTAATGCTAAACAAGAAATTTGAACAAGATGGCGCAAGTCAAAATTGATCACTATCAAT	7029
Qy	282	ValArgGlyGlyGluAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeu	301

Alignment scores:
Pred. No.:

Oy		302	LeutIcGlyaspValPheHisGlyLysTyrGluThrLysLysAsnLeuIcGluAsn	321
			: : :	
Dd	7087	GCTAAATTCATGCATTACAATAATTACTAATGCTACCACTACACAATTAGTTAATGAT	7146	
Oy		322	AlseRlaGluthisGlyTyrPheaspGlyArgTrpLeuAspArgServaspVallie	341
			: : :	

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

xy	342	teuf	loas	pas	nu	ma	as	pr	va	se	le	u	te	u	ta	s	pr	u	u	g	u	u	ta	u	ne	301
									:::	:::									:::	:::						

[illegible]

QY	002	003	004	005	006	007	008	009	010	011	012	013	014	015	016	017	018	019	020	021	022	023	024	025	026	027	028	029	030	031	032	033	034	035	036	037	038	039	040	041	042	043	044	045	046	047	048	049	050	051	052	053	054	055	056	057	058	059	060	061	062	063	064	065	066	067	068	069	070	071	072	073	074	075	076	077	078	079	080	081	082	083	084	085	086	087	088	089	090	091	092	093	094	095	096	097	098	099	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455
----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

DD /243 AAAAAACCCIAIA AIAAATC AAAACACCCAAATGCCAACACAAATAAATAA /23

[illegible]

DD / 254 GAAGCAATATATAAGTTCATCGTCTCAGGAAAGAGCAGATGCTAATATGCTAATAGCG / 333

QY 40Z 1YASNLGUA1AVATA1GATALEUSETASIASPLEUTTEATATATATGTYIPNEASII 421

```
Db 7354 CACACTACTGATGAAGTTAATAATATAATAATCAAA-----GCTGTTCAAAATATTAAAC 7407
Qy 422 MetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSer 441
Db 7408 GCTGTTCAAGTTGAAGTTATC-----AAGAAACAAACCGTTAAACACCAATTAATCAAA 7461
Qy 442 PheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGlu 461
Db 7462 TTCATTGATATCAAAAGAAAATTAATGCA-----AATAGCGCTGATGCAACACTAGAA 7515
Qy 462 ProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro----- 479
Db 7516 GAAAAGCTGAAGTAATAGATTGCTTCAAAATGTACTACTTCCACATCAGATGAATTT 7575
Qy 480 IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuValAlaIleLysAla 499
Db 7576 GCTAATGTAGATCATAACACAGAGGTGATCAAGCTTTTATAGTAAGCTAGACCAAAATC 7635
Qy 500 ArgHisLeuTyrAspMetProAspArgValLeuAlaIleAsnHisAspAspGlyVal 519
Db 7636 GAGGAATTTGACCACAAAGTTAGTAAGAAA----- 7665
Qy 520 AsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeu 539
Db 7666 ---CGAGATGTTTAAATGCAATCCAAAGACATTTAATTCACAAACTCAAGAAATA--- 7719
Qy 540 ProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArg 559
Db 7720 -----CAAGAGAACCAAGAGCTACGAATGAAAGAAAACCTGAAGCATTAATATAA 7770
Qy 560 -----LysThrProAlaAspValTyrGlnSerLysLysVal 571
Db 7771 ATAAACCAATTACTTAATCAGGCTAAAGTAAATATTGATCAAGCACAGTCAAAATAAA--- 7827
Qy 572 ProLeuTyrValPheValAlaSerAspLysProArgaspGlyGlnIleGlyLeuGlyTyr 591
Db 7828 -----GATGTAGATAGTGCAGAAA----- 7845
Qy 592 GlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAsp 611
Db 7846 -----ACACGTAGTATTCAAGATATAGAG---CAATTCACACCACATCCA 7887
Qy 612 GlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeuTyr 631
Db 7888 CAAACAAAGCAACCGCGCTCACAGATTAATGAA-----AAAGCTAACCAACAACAA 7941
Qy 632 AlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGln 651
Db 7942 AGTACT---ATTGCAATCATCCT-----AATTCACAAATTTGAAGAAGA 7983
Qy 652 GlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHis 671
Db 7984 CAGCAA-----GCAAGTGCAAAACTACAAAGAA 8010
Qy 672 GluIleSerArgSerIleIleGlnAsnGlyClyTrpAsnArgThrTyrSerLeuArgTyr 691
Db 8011 GTTCTTAAAAAAGCCATAGCT----- 8031
Qy 692 ArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAsp 711
Db 8032 AAATAGATAAAGTCAAAACCAAT-----GATGATGTAGAAAAGACT 8073
Qy 712 PheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThr 731
Db 8074 GTAGTAAACGGAATAGCTGAAATTTGAAATATATATTACCTGCTACT-----ACA 8121
Qy 732 ValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluVal 751
Db 8122 GTTAAAGATAAAGCTAAGCTGATGTAATGCTGAAAGAGAGCAAGAAACCTACAATTT 8181
Qy 752 GlySerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGly 771
```

```
Db 8182 AATAGTAATGATCAAGCAAGCACTGAAGAAAAAATTAGTTGCT----- 8223
Qy 772 ValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGly 791
Db 8223 ----- 8223
Qy 792 IleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisVal 805
Db 8224 -----AGTGACAATTTAAATCACGTT 8244

RESULT 40
US-09-147-236-1
; Sequence 1, Application US/09147236A
; Patent No. 6316251
; GENERAL INFORMATION:
; APPLICANT: TONOUCHI, Naoto
; APPLICANT: TSUCHIDA, Takayasu
; APPLICANT: YOSHINAGA, Fumihiko
; APPLICANT: TAHARA, Naoki
; APPLICANT: HAYASHI, Takahisa
; TITLE OF INVENTION: NOVEL GENE, GROUP OF GENES, AND NOVEL BETA-GLUCOSIDASE
; FILE REFERENCE: 6537-011-0PCT
; CURRENT APPLICATION NUMBER: US/09/147,236A
; CURRENT FILING DATE: 1999-04-08
; EARLIER APPLICATION NUMBER: PCT/JP97/03633
; EARLIER FILING DATE: 1997-10-09
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 16836
; TYPE: DNA
; ORGANISM: Acetobacter xylinum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (869)..(1891)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3101)..(5368)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (5373)..(7778)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7784)..(11761)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11764)..(12231)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12448)..(14652)
; FEATURE:
; OTHER INFORMATION: n at positions 15741 and 15767 may be a, g, t, or
; OTHER INFORMATION: c
US-09-147-236-1

Alignment Scores:
Pred. No.: 0.0162 Length: 16836
Score: 130.50 Matches: 198
Percent Similarity: 33.82% Conservative: 146
Best Local Similarity: 19.47% Mismatches: 367
Query Match: 2.76% Indels: 307
DB: 4 Gaps: 51

US-09-914-168-2 (1-919) x US-09-147-236-1 (1-16836)
Qy 47 ThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThr---ProGluGln 65
Db 8324 ACCCAGTGGGACCAAGCGCGCGGGCTGCCCGGGATCGTTGGTCCAAACCCGAGAAAT 8383
Qy 66 IleGlnAlaArgLeu-----AsnAlaAlaGlyLeuAsnAlaLysProGlnSerGln 82
Db 8384 TACCGCGCCCAAGTCGCTTTGGCCCAAGCCCTTTGACCTATATAATACCTCGACCCCGCATGAA 8443
```

D	b		9353	GATGTGGCTGAAGCGGCGCGTGATG-----CAGCCCATCTCTGGCCAATCCTGTACC	9406
Q	y		401	AlatyrAsnLeuGlnAla-----ValArgAlaLeuSerAsnAspLeuIleAla	416
			:::		
D	b		9407	CGCAGGACGCCAGCGCGGTATCCTGTATACCTATGCAGTGGCAATGATCGCATGACC	9466
			:::		
Q	y		417	ThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGln	436
		:::	::::		
D	b		9467	CGCCAGCTTCGGCGGTCCTGCGCGCGGATTATCCCCCGCATCCGTTCCATTGCC	9526
			:::		
Q	y		437	AsnAspGlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp	456
		::::	::::		
D	b		9527	--GAGGAATGGAAATCAAGCAGGAACCTGCCACGCCGCTGTCGATGGTATCCAACCCG	9583
			:::		
Q	y		457	GluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAsp	476
			:::		
D	b		9584	GTGCGGTGATCCGCGAGGCCCTTCCAGCGCTATCCGACCGCGCGGTGGGTGGCG	9643
			:::		
Q	y		477	IleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAla	496
			:::		
D	b		9644	GTGCGCACCTGTTCCGCGAGCGTGACATGGTGATGCGCGCATGCCGCTGCGGTATC	9703
			:::		
Q	y		497	AlalysAlaArgHisLeuTyrAspMetProAspAspArgVal-----	510
			:::		
D	b		9704	GCCTCGAGCGCACCATCGACTTTCGCGCCAGCAGCGCCTGTCTACGCCACCGAATAC	9763
			:::		
Q	y		511	LeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAla	530
			:::		
D	b		9764	ATGAAGTACGACACCGCGGTGGCAGCGCAGC-----CTGCTGGCCCCCTGGGGATGCG	9820
			:::		
Q	y		531	Val---SerAlaValAlaArgAlaIleLeuProAspGluSerGlu-----	544
			:::		
D	b		9821	ACGGTTCGGCGCGGCAATCGCTGCTGCCGAGCAGATCGACATTCGACGCACTG	9880
			:::		
Q	y		545	AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAsp	564
			:::		
D	b		9881	CGCATGGGCATCTCGGTGGCGCAGTCCGATGTGCTCAACACGCGTGGCGCAGCGAG	9940
			:::		
Q	y		565	ValtyrGln-----	567
D	b		9941	GCCTACGATCATCTGGCGCGCGCTCGAGCGCCGAGCGGAGGAGACATCGCCCAAGCTG	10000
Q	y		568	-----SerLysLysValProLeuTyrrVal	575
			:::		
D	b		10001	CGCTCGCGCGCTCTATAACGGCCACGCAAGCGGCAAGCGCTCGAGATCGACCTT	10060
			:::		
Q	y		576	PheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAspThr	595
			:::		
D	b		10061	CGGTGTCTGCCACAACCCGACGATCTTGATGC- ¹ CGCGAGCGCGGCTGCGAGCGGC	10119
			:::		
Q	y		596	GlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrrGlnAla	615
			:::		
D	b		10120	GGTCACAGCAACACAAACAGCTCTTGGCCACCG-----TCTCGGATGGATGGCGTGCA	10173
			:::		
Q	y		616	Gly-Ala-----	617
			:::		
D	b		10174	GSMAAGCCGATGGATGCCCGCGCTGCTGGCCATGCGCTAGCTGACCGCGCGATGG	10233
			:::		
Q	y		618	-----GluLeuArgLeuSerGluAspLysLysGlyValLysLe	630
			:::		
D	b		10234	CCATGGTCAGCGCACCATCGAGGACCTCGCGCGCGCTATGACCTGCGCCTGCAGCAGT	10293
			:::		
Q	y		630	utYrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTy	650
			:::		
D	b		10294	CGAGGCGACGCGGCGCGCTGTGTCGCGTGGG-----GCACA	10324
			:::		
Q	y		650	rGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGl	670
			:::		
D	b		10333	TGAAGAAGCGTGTGCCCGCCATCGACCAACCCGTT-----CACTCGCTGGCTACGG	10386
			:::		
Q	y		670	uhISgluIleSerArgSerIleIleGlnAsnGlyIleGlnAsnGlyTrpAsnArgThrSerLeu	690
			:::		

Db 10387 GCATCAGGTGGAACCTGGCGCGCGGTGACCGGTGGC-----TCCTACAGTGCCGA 10437
Qy 690 gTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProVa 710
Db 10438 G-----GCGGCATCGCCCGATAG----- 10456
Qy 710 lAspPheValAsnGlyLysProSerGlnAlaLeuLeuAlaGlyValAla-----Va 728
Db 10457 -----TCGGACACAGATGCTCTCCATTCCTGCGGCAGAT 10491
Qy 728 lHisLysThrValAlaAspAsnLeu-----ValAsnProMetArgGlyTyrArgGl 745
Db 10492 CCAC---ACGCTGGGTGAACCTTGACCCCTCCATTCATGCTGGGCTGGGCTTCCGGTC 10548
Qy 745 nArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAl 765
Db 10549 GCGT---TCGGCGCAGCATGGC---ATGGCGCGCTGACGGAGCGAACATTCCTCCATCGT 10602
Qy 765 a-----ArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTy 780
Db 10603 GGGCGCGCTGCGCTGCAGCGCGGTGCTTCGGCC----- 10636
Qy 780 rGlySerAsnArgAlaHisGlnMetThrGlyIleGlnAlaGlyTyrIleTrpSerAs 800
Db 10637 -----CTGACCTTCGATCATACGCCACCATGATGCTGTCGGG 10674
Qy 800 pAsnPheAsn-----HisValProTyrArgLeuArgPhePheAlaGlyGl 815
Db 10675 CCAGCTCAACACAGCGCTCCGTCTATGATGTCGG---CGTTATGGCACGTTTCATGGCAAC 10731
Qy 815 yAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTy 835
Db 10732 GCAGGCTGCCAACACAGTCGCGCGGCACAGTTCGTGTGGCGGCTTGAT-----TT 10782
Qy 835 rLeuThrGlyGly-----GlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPh 853
Db 10783 CTGAGCGCCCAACCATACCCAGCGCATCGCGGCTGGTGCAGGCGAG----- 10828
Qy 853 eMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPh 873
Db 10829 -----GCCGGGTT 10836
Qy 873 eThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTrp----- 887
Db 10837 TCGCGCGGATGTCAGTTCGGCAATAGC-----TCGGTGGCGGTGATGTCTG 10884
Qy 888 -AlaSerProValGlyGlnValArgValAspValAlaThrGlyValLys 903
Db 10885 CGCCTCGCCCATCGCTTCCCATTCACCAACGCTGCTGGGCGGGTTCGAG 10933

Search completed: May 6, 2003, 03:17:43
Job time : 5587 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 6, 2003, 01:42:26 ; Search time 179 Seconds
(without alignments)
6066.535 Million cell updates/sec

Title: US-09-914-168-2

Perfect score: 4727

Sequence: 1 MSKPVLFPANRSMFVALAAY.....TGVKEGNPIKLHFFICTPF 919

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09914168/runat_28042003_151617_6886/app_query.fasta_1.1095
-DB=PublishedApplications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Diosum52
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=40 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09914168_@cgn_1_1_99_@runat_28042003_151617_6886
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAPP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PublishedApplications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US05_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match %	Length DB	ID	Description
1	238.5	5.0	6617	9	US-10-210-296-1
2	140	3.0	17310	7	US-08-781-986A-23
3	139	2.9	5245	9	US-10-080-505-12
4	137.5	2.9	7104	10	US-09-815-242-4580
					Sequence 4580, Ap

5	137.5	2.9	7107	10	US-09-815-242-8291	Sequence 8291, Ap
6	136.5	2.9	38584	9	US-10-114-170-50	Sequence 50, Appl
7	136	2.9	8673	10	US-09-815-242-7474	Sequence 7474, Ap
8	136	2.9	12404	9	US-10-108-605-236	Sequence 236, App
9	135	2.9	5629	9	US-10-114-170-243	Sequence 243, App
10	135	2.9	6228	10	US-09-815-242-4760	Sequence 4760, Ap
11	135	2.9	6561	10	US-09-815-242-8815	Sequence 8815, Ap
12	134	2.8	3121	10	US-09-765-272-117	Sequence 117, App
13	133.5	2.8	14536	9	US-10-160-758-6	Sequence 6, Appl
14	133.5	2.8	14536	9	US-10-160-758-7	Sequence 7, Appl
15	133	2.8	17388	10	US-09-815-242-8512	Sequence 8512, Ap
16	132.5	2.8	1488	10	US-09-815-242-9166	Sequence 9166, Ap
17	132.5	2.8	2571	10	US-09-815-242-7037	Sequence 7037, Ap
18	132	2.8	7434	10	US-09-815-242-4761	Sequence 4761, Ap
19	132	2.8	7437	10	US-09-815-242-8869	Sequence 8869, Ap
20	131.5	2.8	6077	10	US-09-815-242-4648	Sequence 4648, Ap
21	131.5	2.8	6077	10	US-09-815-242-8513	Sequence 8513, Ap
22	131.5	2.8	31096	7	US-08-781-986A-59	Sequence 59, Appl
23	131	2.8	7302	10	US-09-815-242-4780	Sequence 4780, Ap
24	131	2.8	18846	10	US-09-815-242-8898	Sequence 8898, Ap
25	130.5	2.8	3552	10	US-09-815-242-4174	Sequence 4174, Ap
26	130.5	2.8	3567	10	US-09-815-242-8027	Sequence 8027, Ap
27	130.5	2.8	7311	10	US-09-815-242-4779	Sequence 4779, Ap
28	128.5	2.7	3336	10	US-09-954-456-1807	Sequence 1807, Ap
29	127.5	2.7	5832	10	US-09-881-752A-307	Sequence 307, App
30	127.5	2.7	8673	10	US-09-815-242-7313	Sequence 7313, Ap
31	127	2.7	2545	10	US-09-974-300-1524	Sequence 1524, Ap
32	126.5	2.7	11185	10	US-09-880-107-3311	Sequence 3311, Ap
33	124.5	2.6	9534	9	US-09-954-531-1380	Sequence 1380, Ap
34	124.5	2.6	10302	10	US-09-782-378A-23	Sequence 23, Appl
35	124.5	2.6	14155	9	US-10-108-605-102	Sequence 102, App
36	124.5	2.6	50937	9	US-09-808-880-1	Sequence 1, Appl
37	124	2.6	1815	9	US-10-107-649-1	Sequence 1, Appl
38	124	2.6	5346	10	US-09-738-877-2	Sequence 2, Appl
39	124	2.6	6608	10	US-09-738-877-1	Sequence 1, Appl
40	124	2.6	6608	10	US-09-880-107-3439	Sequence 3439, Ap
41	123.5	2.6	3349	9	US-09-952-267-2	Sequence 2, Appl
42	123.5	2.6	3753	10	US-09-801-368-363	Sequence 363, App
43	122.5	2.6	2609	10	US-09-789-561-48	Sequence 48, Appl
44	122.5	2.6	6026	9	US-09-995-749A-1	Sequence 1, Appl
45	121.5	2.6	2556	9	US-09-738-626-3499	Sequence 3499, Ap

ALIGNMENTS

RESULT 1
US-10-210-296-1
; Sequence 1, Application US/10210296
; Publication No. US20030021802A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Products Inc.
; TITLE OF INVENTION: LAWSONIA INTRACELLULARIS PROTEINS, AND RELATED
; TITLE OF INVENTION: METHODS
; TITLE OF INVENTION: AND MATERIALS
; FILE REFERENCE: PC10589A
; CURRENT APPLICATION NUMBER: US/10/210,296
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US/09/689,065
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ IDS NOS: 102
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 6617
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-10-210-296-1

Alignment Scores:
Pred. No.: 1.86e-13 Length: 6617
Score: 238.50 Matches: 198
Percent Similarity: 34.72% Conservative: 137
Best Local Similarity: 20.52% Mismatches: 331
Query Match: 5.05% Indels: 239

DB: 9 Gaps: 38
US-09-914-168-2 (1-919) x US-10-210-296-1 (1-6617)
QY 50 AsnGlnAlaLysAlaGlyAsnProValLeuLeuThrProGluGlnIleGlnAlaArg 69
DB 3764 AATGCTGCTCAAAAGACGATCCTTCTATTGTGGTCTCCCATTTCAAAATTAATGGCTCA 3823
QY 70 LeuAsnAlaAlaGlyLeuAsnAlaLys---ProGlnSerGlnAlaLeuAspValValAsn 88
DB 3824 TCAAATGATGAAGAGTTACAAACAGAACTACCAATGCTTCTTGCAACTGCATTAAGAAT 3893
QY 89 PheAspGlnSerProIleSerArgIleGlyGluGlnSerProLeuGlyLeuAsp 108
DB 3884 -----AAGGGA 3889
QY 109 MetSerValIleGluThrThrProLeuSerLeuGluLeuPheAlaGlnGluSer 128
DB 3890 TTTTCGTGTCATCCCTAATAAATCTGCATTAATCTT-----CTATATAACAAATAATC 3943
QY 129 ThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnProAsn 148
DB 3944 TCCCAACTTAATATTCTTACTGCAAAAAGGTA-----GCTCAACAACCTCCAT 3991
QY 149 SerGluValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeu 168
DB 3992 GCTGACTATGATGA----- 4006
QY 169 TyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyr 188
DB 4007 TACGGCATTTCATCAACACAGGTGAAT-----TTTAGT 4042
QY 189 GlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSer-----HisGlnLys 205
DB 4043 ATTGATAGTAGGCTTATTGATAGTACAGGTGTAGCATCTGCACGCTCCATTAATACAGAA 4102
QY 206 ThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAlaMet 225
DB 4103 AAACCAAAATTTAATAGCTAAATATTGCTGTAAACAGAACTGTCTGAAGCTATTAAGT--- 4159
QY 226 AspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaVal 245
DB 4160 -----AATGGCTTATAAAGAAA---AACACTATTGCTGATGATACGTATTATCATGGCCTT 4210
QY 246 GlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIle 265
DB 4211 AAAGTTCTTGATCCTGATGAATCTTACACGACTACTATTATAAAGGGAGATCATACT 4270
QY 266 IleHis-----AspLeuGlyGluProValTyr 274
DB 4271 GATCATGCCAAAATTAATGCAGAAATCAAAAATAATATGGGAATTAGGA-----TATTTT 4324
QY 275 IleAspTyrArgAla---ValGluValArgGlyGluGlyAlaAspAspLysAlaPheThr 293
DB 4325 AGTGATGCTCTGCAAGTATTGAAGAAAGCGGGAAGGA-----CGATTACTTTGTA 4375
QY 294 ThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisGlyLysTyrGlu 313
DB 4376 TTTACTGTACAAGAAAGCCCTAAATTTACAGATGTTTGTGTTCAAGGCTCAAAAGCTGTA 4435
QY 314 ThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyr 333
DB 4436 AGTATCATCAATCTTCTGTCGAATCAGTTCTAAAAAGGA----- 4477
QY 334 LeuAspArgSerVal-----AspValIleLeuProAspAsnThrAlaAspValSerLeu 351
DB 4478 -----TCAGTTATTAGTAGACTATTGTCCTCAAGATATATCAAAAATAATACCGAC 4528
QY 352 IleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIleAspPro 371
DB 4529 CTCCTATAGAAAGAGGCTACTATCTCGCTGAA-----GTTAATTATGAATAAAGAG 4582
QY 372 LysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeuGlu 391

DB 4583 AAAGAAAATACTTCTTCTGCA----- 4606
QY 392 GlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSer 411
DB 4607 CTATTGTTAAACAGTAATGAAGGGAATACTTTATATTAAGAGATGTCGGAATGAAGGA 4666
QY 412 AsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGlu 431
DB 4667 CTTCGAACAATAAAGCTAAA-----ACTTTAAAAAAGAGCTTAGCATTACAGAA 4717
QY 432 Arg-----GluGlnIleGlnAsn 437
DB 4718 CGTAATTTTTCATGCTTACTGGAACAGGTGTATTACGTGAAGAATACTTTCAACCT 4777
QY 438 AspGlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGlu 457
DB 4778 GACTCTATAGCAATC-----TCTGCCCTATGCCATGAATCATGGCTATGTAGATATT 4828
QY 458 SerThrLeuGluPro-----ValIleGluThrValIleGluLeu 469
DB 4829 CAAGTTGCTTCCACCTGAAGTAACATTCAATCAAAAAGGAATTGTTATTACATTTAGAGTA 4888
QY 470 ThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGln 489
DB 4889 AAGAAGGTAAAGCCCTATATAAATAAGAAAAATAGACTTTTAAAGGAGATCTTTATTTCAGACA 4948
QY 490 AspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAspArg 509
DB 4949 AATGAACAACCTCTTAAGGTAAACAAAATTTCATCATATAAAAATATATGACCACTATTTT 5008
QY 510 ValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAsp 529
DB 5009 TCTCTTCTGTATGCAAGATGAT-----GTAAAAGCATTAAACAGATTTTATTATTTCAGAT 5062
QY 530 AlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu----- 546
DB 5063 TATGGTTATGCATTTGCTGAAGTAGATCTT---CAAAACAACCAAAAATGAAGAAGATGCA 5119
QY 547 -----ValIleAspLeuProGluArgThrAlaLeu----- 556
DB 5120 ACAATTGATGTTACTTCTTATTGATAAAAACAACAAAGTCTTTCTTCGTAGATAAATT 5179
QY 557 -----AlaAsnArgLysThrProAlaAspValTyr----- 566
DB 5180 GTTGAAGGAAATACTCGTACTAGAGATAATGTTATCCTCGTGAATTACGCGCTTGCCTGAT 5239
QY 566 ----- 566
DB 5240 GGAGATCTTTTAAATGTCACATCTCCGAGGCTCTAATGAATCCCTTAACCGCCTTGGC 5299
QY 567 -----GlnSerLysLysValProLeu 573
DB 5300 TATTTTAAACCAAGTAGATACAGATACACTGCCTACAGGAAAGATGATGAAGTTGATCTA 5359
QY 574 TyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrGlySer 593
DB 5360 CTTGTAAAGTT-----CAAGAGCTCGAACAGGTGCAATCACAGGTGGTGTGGTTAC 5413
QY 594 AspThrGlyThrArgLeuValThrLys-----PheGluHisAsnLeuIleAsnArg 610
DB 5414 TCACACATTTCTAAATTTGGTGTTCAGGAAGTATCTCAGAAAGAAACCTTATCGGGGAAA 5473
QY 611 AspGlyTyrGlnAlaGlyAlaLeuArgLeuSerGluAspLysLysGlyValLysLeu 630
DB 5474 ---GGTTATATTTAAAGTATTGAAGGTTTATTCTTAGTAGTCAATCTTCTCTTGATCTT 5530
QY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
DB 5531 TCTTTTACCATAATCCTCGTGTGTATGATACAGAC----- 5563
QY 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeu----- 664

```
Db 5564 -----TTTGGCTTTAGTAATAACATTTATACGCTACGAGATGAATGGGATGAC 5611
Qy 665 -----SerThrArgThrLeuGluHisGluIleSerArgSerIle 677
Db 5612 TTCGTAATAAAACCTTATGGAGATACCATAGCTCTATTTCACCTATATA----- 5659
Qy 678 IleGlnAsnGlyIleGlyTyrPheAsnArgThrTyrSerLeuArgGlyTyrArgLeuAspGlyLeuLys 697
Db 5660 -----GCAGATAATTCATCTATCTTTGCTGCATATCGAATGATCAATATCGT 5707
Qy 698 ThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysPro 717
Db 5708 CTATATGATATTCATCTACACGACCAAGCTCTTATCTTGACTAT---CAAGGGAATAAT 5764
Qy 718 SerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuVal 737
Db 5765 ATTTCTAGTGTAGTAAGTGGTGTATTTACTTTTGTATCTACA-----GACAGTCGTCAG 5818
Qy 738 AsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuVal 757
Db 5819 AGACCATCTAAAGGCATATTGCAAACTAATCTTGAATATGAGAGTGGTGGTCTTGGT 5878
Qy 758 SerAspAlaAsnMetAlaIleAlaArgNlaGlyIleSerGlyValTyrSerPheGlyAsp 777
Db 5879 GGTAATGATAACTTCTTCAAGCCCAATTCGCTCAACTACAAAGGATTTTACTCAATT----- 5932
Qy 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIle 797
Db 5933 -----TCAGAGTAAACCATATATATACATTCGCTACACGTCGAGGTCGACGT 5983
Qy 798 TrpSerAspAsnPheAsnHisValProTyrArgLeuArgPheAlaGlyGlyAspGln 817
Db 5984 TATAAGATATGATAAAACCTGTGCACATATTGACCATTTTATTTGTTGGTATAGAT 6043
Qy 818 SerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLys---GlyTyrLeu 836
Db 6044 AGTATTAGAGGATGATGACAGAGATCTTGACCAAAAGATCCTCGCTTTGGAGATGAA 6103
Qy 837 ThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAsp 856
Db 6104 ATTGCTGTGTAGTAGTGCCTTTCTTAACTAGATATATTGGACATTTCCAGCCAGAG 6163
Qy 857 LeuArgLeuAlaVal-----PheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThr 874
Db 6164 CTAGTCTTGCATGATTGCTTCCATCTATGACATAGGATTCACAAACAGATTCGTACAAACT 6223
Qy 875 AsnAsp-----ThrLysIleGlyAlaGlyValGlyValArgTyrAlaSerPro 890
Db 6224 TCTAACCCATCTCTAAACTCAACAAATCATATGCGCTTGACATTCGCTGGCTTCACCA 6283
Qy 891 ValGlyGlnValArg 895
Db 6284 ATGGGAGATTTGCGA 6298
```

RESULT 2

```
US-08-781-986A-23/c
; Sequence 23, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; STREET: Human Genome Sciences, Inc.
; ADDRESS: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
```

```
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
FILING DATE: US/08/781,986A
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248PP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 17310 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-23
Alignment Scores:
Pred. No.: 0.00607 Length: 17310
Score: 140.00 Matches: 177
Percent Similarity: 32.16% Conservative: 124
Best Local Similarity: 18.91% Mismatches: 276
Query Match: 2.96% Indels: 359
DB: 7 Gaps: 45
```

US-09-914-168-2 (1-919) x US-08-781-986A-23 (1-17310)

```
Qy 205 LysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAla 224
Db 12111 AAAGTAAATAAATAACAACTTATAAGAACCAATTCGCGTATTTTGCACATATGCGT 12052
Qy 225 MetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAla 244
Db 12051 CAAAGTGCAGAT-----AATCAAGCGCAGTGCACAAACATATGGCG 12010
Qy 245 ValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspVal 264
Db 12009 CAATATGCTGTCGATCATGATAGCATTCATCCACGGCA----- 11968
Qy 265 IleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGly 284
Db 11967 -----CGTTATTGCAAGTAACACAATAGACTAC----- 11938
Qy 285 GluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGly 304
Db 11937 -----GTAGAGGATCGT-----TATATCGTTTCAGAACAACTCTTTATACTTAGT 11893
Qy 305 AspValPheHisHisGlyLysTyrGluThrLysLysAsnLeu----- 318
Db 11892 CATCCATTTTCAT-----CCGACTCTAGAGTGCAAGTGGTTCAGAACGA 11845
Qy 319 ---IleGluAsnAlaSerAlaGlu---HisGlyTyrPheAspGlyArgTrpLeuAspArg 336
Db 11844 GATTTAGAGAAATATGACCCGCAATGTCATACATCATTCCAATTCATTTAGTGTG 11785
Qy 337 SerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGly 356
Db 11784 CATCAAGATGTTCTGCTCAGCGC-----TATGTAGAAGGT 11749
Qy 357 ThrGlnTyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGln--- 375
Db 11748 AAAGAAGATCAGGTTGAGAAAGTGTGTCAATTTAGCAGACATATATCAGAGATA 11689
Qy 376 -----LeuThrThrAspProAspLysLeuProValLysArg----- 387
Db 11688 CCCAAGATTTTATTTTATTATCCACACATCTCTTATCAANTCAATCTGTGTGCGACAGCAT 11629
```

QY 388 -----GluLeuLeuGluGlnLeuLeuThrValAsnMetGly----- 399
Db 11628 CCACAGTATATGCAATATAGTGAACAGGTTTAAATAAAGACCTTGGCGTGTAT 11569
QY 400 GluAlaTyrAsnLeuGlnAlaValArgAlaLeu----- 410
Db 11568 TCAGTGTACCGGACGCTTTCGGTTAGAACTGATTTTCAAAAGCATTAACATTTATTTA 11509
QY 411 -----SerAsnAspLeu----- 414
Db 11508 AAATTACCAGTACACGCTTAAATCACTAAATTTTATACGTACGAATGACCTTGAACAGATT 11449
QY 415 -----IleAlaThrArg 418
Db 11448 GAAGGACAATTTGATCGCGCGCAAGTTATCGCATCAGTCAAAGATGAGGTTGAACAACCC 11389
QY 419 TyrPheAsnMetValAsnThrGlu----- 426
Db 11388 CATTTTAAATTTGATGTTTGAAGAAGGATATCGTGCATTGTTACCGAATCCATTAGGGCAA 11329
QY 427 IleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSer--- 445
Db 11328 ACAGTTGAACCTGAATGATTTATTAACAATAGTGCCATGATTTCTGCGNAGGGATA 11269
QY 445 ----- 445
Db 11268 CGAATTACCAGTCTGATAAAGATATTCATGATTGGCGTCATTTATTGAAACGATGCCT 11209
QY 446 SerSerSerArgThrGluProAlaGlnValAspGluSer----- 458
Db 11208 GATTCACGATGCTCAAGTTATCACAAAGTATGAGCAAAAGTGGTTTAGCACCAGAAGCA 11149
QY 459 -----ThrLeuGluProValIleGlu----- 465
Db 11148 TGGCTTGAATGTTATTGATCGTACATTTACCCGATATTAAAGCTGTTAGTAACACA 11089
QY 466 -----ThrValGluLeuThrAspGlyIle----- 473
Db 11088 GGCATTAGTCTAGAAGCACATGTACAAAATACATTAATTTGAATTAAGAGATGGCATACCC 11029
QY 474 -----LeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGln 489
Db 11028 GACGTATGCTTGTGACAGATCTTGAAGCATTTGTCTATCTAGACGATTCGTACTGAA 10969
QY 490 AspLysLeu-----AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAsp 507
Db 10968 AAACAGCTTGTGCGAAATGTGTGGCAGCATCAAGCCCTGTTCTATAT----- 10921
QY 508 AspArgValLeuAlaIleAsnHisAspAspGlyValAsnArg-----SerIle 523
Db 10920 -----GCACATGATGAAGCATGCGTCGTCTAAATATTACGTTGTA 10879
QY 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543
Db 10878 GTAAATCACTTAGCATTATTAGTATCACTATTGGTAAAGCG-----ACT 10834
QY 544 GluAsnGluValIle-----AspLeuProGluArgThrAlaLeuAlaAsnArgLysThr 561
Db 10833 AGAAATGAAGTTGTGTTATGGCAACTTGTAGGCATCGTCTTATGACTTGGAAAAAAGAA 10774
QY 562 ProAla-----AspValTyrGlnSerLysLysVal 571
Db 10773 TAGCGGAATAACGCGCATTTTGTGTTAGTGTGTAGAAATTTATATCAACGCGGACCAT 10714
QY 572 -----ProLeuTyrVal 575
Db 10713 GCGGCTAAACGGAATTTGATGAGTAATTAATGATTTGGTGCAACCCCTATTATTATACA 10654
QY 576 PheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTrpGly----- 592
Db 10653 CATATACCAATCCAAATTTGTGCATACAAAGGAGGTATCTGTTGTGAATCAACAAATTC 10594
QY 593 -----SerAspThrGlyThrArgLeuValThrLys-PheGluHisAs 606

Db 10593 TAATCGTGTAAAGACTAGAGTGTACCACTGGTATCATCATCTTATTATGAGAATAT 10534
QY 606 nLeuIleAsnArgAspGlyTyrGlnAlaGly----- 616
Db 10533 TGTGTGTATAAAGCGTCATATCAAGCGGTGTCGGTCACTTTTACAAATAGAAGCACATGA 10474
QY 617 -AlaGluLeuArgLeuSerGluAspLysLysGlyVal-----LysLeuTyrAlaThrLys 634
Db 10473 TTCAGAGTATCGTTTACTGCTGAAAGACACATACGTTTGTATGATATGATATCATATC 10414
QY 634 sProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGln----- 652
Db 10413 ACCAATTGAGCGTGTCTAGGAGATCAGCAGATACAACAACAGACTATACAAATATT 10354
QY 653 -----GluValPheGlyHisSerThrAsnGlyPheAspLeuSer-----ThrAr 667
Db 10353 GAGAGAGGTGTTGATTTTACATTTCTTAAAAATGATGAAAGACTAGAACAAATTTATTCGA 10294
QY 667 gThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThrTy 687
Db 10293 GTTATTACAGACAGATTAT-----AAAGATACGCA 10264
QY 687 rSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGlnAs 707
Db 10263 AAGTATGCAGTATCGA-----GAATCAAAACCCACGAGCAACCTTGAGACATTTAACGA 10210
QY 707 pLeuProVal----- 710
Db 10209 CTATGAATTTTATGCCATCGAAGGCATCAGTATCATCCTCAAGTTACAAATACCTTAGG 10150
QY 711 -----AspPheValAsnGlyLysProSerG1 719
Db 10149 ATTTACGTTGAGTGATAATTTGAAATTTTGTCTCTGATTTTGTACCAACCTTAACTGCA 10090
QY 719 nGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsn---LeuValAs 738
Db 10089 GTGGTTAGTATCGCAAAAGATAAAGTAAAGACGCGGTATCAAGAAATGTTGTAGTTAA 10030
QY 738 nProMet-----ArgGlyTyrArgGlnArgTyrSerLeuGluVal 751
Db 10029 CGNAATGTTACGTCACAAAGTTGGCGCATTAAGACTTAT---GAACATTTTGTACAGCAAT 9973
QY 751 lGlySerSerGlyLeu---ValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSe 770
Db 9972 TGAAGCATCTGGCAACACATGTAATGATGTGAGATGATACCTGTACACCCA----- 9921
QY 770 rGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrG1 790
Db 9920 -----TGGCAGTTT---GAACATGTCATCCAAAGTTGATTTGGCTGAAGAAAGCGTTAA 9871
QY 790 yGlyIleGlnAlaGlyTyrIleTrp-----SerAspAsnPheAsnHisValProTy 807
Db 9870 TGGCACA-----GTACTATGTTAGGGAAAGTGTAGCTATATATC----- 9828
QY 807 rArgLeuArgPhePheAlaGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLe 827
Db 9827 -----CCTCAACAATCGATTCTGT-----ACGAT 9805
QY 827 uSerProIleSer---AspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThr 846
Db 9804 GTCGCAATAGACACGACAAA-----ATATTATTAAAGGTACC 9766
QY 847 AlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGly 866
Db 9765 ATAAGTATACGAACACTTC----- 9745
QY 867 AsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArg 886
Db 9744 -----AACGAAACGAGTGTGGCCCTCATACAAATTCGAAA 9709
QY 887 TrpAlaSerProValGlyGlnValArgValAspValAlaThr 900

Db 9708 TGCAGCGCAATTACGGATTGGTTAAAGCAGATACAGCAACA 9667

RESULT 3

US-10-080-505-12

; Sequence 12, Application US/10080505

; Publication No. US20030073166A1

; GENERAL INFORMATION:

; APPLICANT: St. Gene, Joseph W.

; TITLE OF INVENTION: HAEMOPHILUS ADHERENCE AND PENETRATION PROTEINS

; FILE REFERENCE: A-5941-1/RET/DCF/DHR

; CURRENT APPLICATION NUMBER: US/10/080.505

; CURRENT FILING DATE: 2002-02-22

; PRIOR APPLICATION NUMBER: US 08/296,791

; PRIOR FILING DATE: 1994-10-25

; PRIOR APPLICATION NUMBER: US 09/839,996

; PRIOR FILING DATE: 2001-04-20

; NUMBER OF SEQ ID NOS: 58

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 12

; LENGTH: 5245

; TYPE: DNA

; ORGANISM: Haemophilus influenzae

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (430)..(4740)

; OTHER INFORMATION:

US-10-080-505-12

Alignment Scores:

Pred. No.:	0.00133	Length:	5245
Score:	139.00	Matches:	157
Percent Similarity:	34.10%	Conservative:	153
Best Local Similarity:	17.27%	Mismatches:	352
Query Match:	2.94%	Indels:	248
DB:	9	Gaps:	36

US-09-914-168-2 (1-919) x US-10-080-505-12 (1-5245)

Qy	13	MetProValAlaLeuAlaAlaTyrLeuProLeuMetThrSerGlnAlaLeuAlaGlnGln	32
Db	2384	ATGAGATTGGATCACCAGCATTTAAAGCTGAAAGCTTCCAAATTAAGCGGGAAGTG	2443
Qy	33	AsnAsnProAlaAsnIleAsnHisValProAlaHisAspThrAlaIleAsnGlnAla	52
Db	2444	CGGTG-CTTTCGCAATGTTCTTCAATTGAGGGAATTCGACAGTCAGCAATAATGCA	2502
Qy	53	LysAlaGlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaArgLeuAsnAla	72
Db	2503	AATCCCGCATTTGGTGTGGCCAAATCAGCAAAATACCAATTCGACCGTTCAGATTGG	2562
Qy	73	AlaGlyLeu-----	78
Db	2563	ACAGGATTAACGACTTGTAACACTGTGGATTAAACCGATACAAAGTTATTATTCATA	2622
Qy	79	ProGlnSerGlnAlaLeuAspValValAsnPheAspAspGlnSerProIleSerArgIle	98
Db	2623	CCGACACACAAATTAATGCGCTCTATTAAATTAATGATAATGCAACAGTGAATATTTCAT	2682
Qy	99	GlyGluGlnSerProProLeuGlyLeuAspMetSerValIleGlu-----	113
Db	2683	GGT-----TTAGCAAACTTAATGGTAATGTCACCTTTAATAAATCATPAGCAATTTACA	2736
Qy	114	-----GluThrProLeuSerLeuGlu-----	121
Db	2737	TTGAGCAACATGCCACCAACAGGCAATATCCAACCTTCAAAATCAGCAAAATGCACAG	2796
Qy	122	-----GluLeuPheAlaGlnGluSerThrGluMetGlyIle	133
Db	2797	GTGGACAATGCAAAATTTGAACGGTAATGTGCATTTAAACGGATTCTGCTCAATTTCTTTA	2856
Qy	134	AsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnProAsnSerGluValVal	153
Db	2857	AAAAACAGCCATTTTTCGACCAAAATTCAGGCGGACAAAGACAAACAGTGCACGTGGAA	2916

Qy	154	ProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAlaArg-----	171
Db	2917	AATCGACTTGGACAATGCCTAGCGATCCACATTCGAGATTTAACGCTAAATATAGT	2976
Qy	172	-----LeuPheAsnAspGlyValAsnLysValProArgLeuLysAla	185
Db	2977	ACTGTTAGCTTAAATTCAGCTTATTCAGCTAGCTCAAAATATGCGCCACGT	3027
Qy	186	LysPheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLys	205
Db	3028	-----CACCGCGCTTCATTAGACAGGAAACACCGCAACATCGGCAGACAT	3075
Qy	206	ThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGlnSerAlaMet	225
Db	3076	-----CGTTTCAACACATTG	3090
Qy	226	AspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaVal	245
Db	3091	ACAGTAAATGGTAAATTTGAGCGGG---CAAGGCACATTTCAATTTACTTCTTTATTT	3147
Qy	246	GlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIle	265
Db	3148	GGCTATAAAGCGATAAATTTAAATTTATCCATGACGCTGAGGCGGATTTACACATTATCT	3207
Qy	266	IleHisAspLeuGly---GluProValTyrIleAspTyrArgAlaValGluValArgGly	284
Db	3208	GTTCCGAACACAGCGCAAGAACCCGAGCCCTTGAGCAATTAACCTTGTGTT-----	3258
Qy	285	GluGlyAlaAspAspLys-----	298
Db	3259	GAAGCAAAAGATAATAAACCGCTTATCAGACAACTCAAAATTTACTTTAGAAATGACCC	3318
Qy	299	Val-----ProLeuLeuIleGlyAspValPheHisHisGlyLysTyr-----	312
Db	3319	GTTGATGCGAGTGCATTACGTTATATAATTAGTAGAAGATAATAGCGCAATTCGCTTGCAT	3378
Qy	313	-----GluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHis	326
Db	3379	AACCAATATAAGACAGCAAGTAATGGCAATGTTTAGTAAGACAGACAGCAAGCAAG	3435
Qy	327	GlyTyrPheAspGlyArgTrpLeuAspArgSerValAspValIleLeuProAspAsnThr	346
Db	3436	-----CGAACATAGAACGCCAACAAAGTTGAACAGACT	3468
Qy	347	AlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhe	366
Db	3469	GCTGAACACAA-----ACAAGTAATGCAAGAGTGGCGTCAAAAAGACCGGTGTT	3519
Qy	367	PheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLys	386
Db	3520	TCTGATACCTCGCTGATCAAAAGCCAGTTAGAGGTA-----TTACAA	3561
Qy	387	ArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAla	406
Db	3562	CCCAACAAAGTTGAACCGAGCTGCTGAAAAACAAAAATAAGGCAAAAAAGCGGTCA	3621
Qy	407	ValArgAlaLeuSerAsnAspLeuIleAlaThrArg---TyrPheAsnMetValAsnThr	425
Db	3622	AAAAGACGGGTGTTTTCGATACCTCGCTGATCAAAAGCCAGTTAGACGTATTACAAGCC	3681
Qy	426	GluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSer	445
Db	3682	GAACAAAGTTGAACCGAGCTGCTGAAAAACAAAAATAAGGCAAAAAAGCGGTCA	3741
Qy	446	SerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGlu	465
Db	3742	AGAGCCGCGAGA-----	3753
Qy	466	ThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSer	485
Db	3754	-----GAGTTTTCGAT-----ACCCCGCTTGCATGACTAAGCCGGTTA	3789

QY 486 AsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMet 505
Db 3790 AAGTATTAGAACTCAAACTTGAAGTTATTAAATGCCAA--CAGCAAGTGAAAAAGAA 3846
QY 506 ProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArg-----Ser 522
Db 3847 CCTCAAGATCAAGAGAACCAACCAACCAAAAGACTTGATCAGCGCTTATTCAAATAGT 3906
QY 523 IleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGlu 542
Db 3907 GCGTTATCAGAATTCTCGCAACAGTAATAGATGCTTCTT-----GTTCAAGATGAA 3960
QY 543 SerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrPro 562
Db 3961 TTAGATCGTCTTTTGTGATGATCAAGACAATCTGCCGTGGCAAAAT-----4008
QY 563 AlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysPro 582
Db 4009 -----ATCGCACAGGATAAAAGA 4026
QY 583 Arg---AspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuValThr 601
Db 4027 CGCTATGATCTGATCGTTCGCTGCTATCAGCAGAAACGAACCTTACGT-----4077
QY 602 LysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeu 621
Db 4078 -----CAAAATGGGTGCCAAAGCCCTTA 4101
QY 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
Db 4102 GCTAATCGAGCAATGGGCA--GTTTTCTCGCATACCGCTCAGATAATCTTTTGAT 4158
QY 642 AspGlnLeuArg-----AlaThrLeu-----GlyTyrGlnGlnGluValPhe 655
Db 4159 GAACAGGTAAATCAACGCGACATTAACGATGATGCGGTTTGGCCCAATATCAATGG 4218
QY 656 GlyHisSerThrAsnGlyPheAsp-----LeuSerThrArgThrLeuGluHis 671
Db 4219 GCGGATTTACAAATTTGTTGAACGCTGGGAACGGGAATGAGTGGAGTAAATGGCTGAA 4278
QY 672 GlutIleSerArgSerIleIleGln-----AsnGlyGlyTyrPasnArgThrTyrSer 688
Db 4279 GAACAAGCCGAAAAATTCATCGAAAGCGATAAATATGCGGTGAATGCAAGTTATCAG 4338
QY 689 LeuArg-----690
Db 4339 TTCGGTTAGGCAATGGGCATTGAGCTTATTTGGAGTTAATCGCTATTTATTGAA 4398
QY 691 -----TyrArgLeuAspLysLeuThrGlnAlaProProGluThrTyrGlnAsp 707
Db 4399 CGTCAAAATTTATCAATCTGAGGAAGTGAAGTGAAGACGCTAGCCCTTGCAATTAATCGC 4458
QY 708 -----LeuProValAspPheValAsnGlyLys 716
Db 4459 TATAATCGTGCATTCGAGTTGATTATACATTTACTCCGACAGATAATATCAGCGTTAAG 4518
QY 717 ProSerGlnGluAlaLeuLeuAlaGlyValAla-----ValHis 729
Db 4519 CCTATTCTTCGTCGAATTTATGTTGATGTTTCAAAACGCTAACCGTACAAACCGCTAAT 4578
QY 730 LysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeu 749
Db 4579 AGCAGGFG-----TTCGCAACACCATTTGGACGCTTATTTGGCAAAA-----4620
QY 750 GluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIle 769
Db 4621 GAAGTGGGATTAAGCGGAATTTTACATTTCCAACTTCT-----4662
QY 770 SerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThr 789
Db 4663 -----GCTTTTATTCTAAATCTCAAGGTTTCGCAACTCGGCAACAGCAAAAT 4710
QY 790 GlyGlyIleGlnAlaGlyTyrIleTyr 798

Db 4711 GTGGGGCTGAATTTGGGTATCGTTGG 4737
RESULT 4
US-09-815-242-4580
; Sequence 4580, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR APPLICATION NUMBER: 2001-03-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4580
; LENGTH: 7104
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4580
Alignment Scores:
Pred. No.: 0.00294 Length: 7104
Score: 137.50 Matches: 204
Percent Similarity: 29.88% Conservative: 142
Best Local Similarity: 17.62% Mismatches: 425
Query Match: 2.91% Indels: 387
DB: 10 Gaps: 47
US-09-914-168-2 (1-919) x US-09-815-242-4580 (1-7104)
QY 2 SerLysProValLeuPheAlaAsnArgSerPheMetProValAlaLeuAlaAlaTyrLeu 21
Db 1411 ACTAAGAAGTATTATCCAGATAAATCT-----TTA 1443
QY 22 ProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIle-----38
Db 1444 AAATTATCATATAAGTTAATGTTCCGAATATCGACACACCTAAATAATATTGATTTAAT 1503
QY 39 -----IleAsnHisValProAlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyAsn 56
Db 1504 GAAATAATTAACATATCGTACTGCTTCAGATTTGTAATTAATATGCGCAA-----1554
QY 57 ProProValLeuLeuThr-----ProGluGlnIleGlnAlaArgLeuAsnAlaAlaGly 74
Db 1555 CCAGAAGTTACACTAAGTGCAGATCCATTTTCAGTAGCAGTTGAAATGAACAAGATGCG 1614
QY 75 LeuAsnAlaLysProGlnSerGlnAla-----83
Db 1615 TTGCAACAACAAGTAACACTCAAGTTGATTAATAGTCATTACACAACAGCATCAATTGCA 1674

Qy	84	----	Leu	84
Db	1675	GAATACAATAAACTTAAACAACAGCAGACAATATTTTAAATGAAGATGCGAATCATGTT	---	1734
Qy	85	AspValValAsnPheAspAspGlnSerProfile	---	95
Db	1735	GAACACTGCAAAATCGTCATCTCAAGCGGCTATTGTAGTGTTTAGTAACTAAATTTACAAGCT	---	1794
Qy	96	SerArgIleGlyGlnSerProProLeuGlyLeuAspMetSerValIleGluGluThr	---	115
Db	1795	GCATTAATTGTAATAACACCACTTCTGTAATTAGATGCTAAAGCTCAAGAAAGGTT	---	1854
Qy	116	ThrProLeuSerLeuGluLeuPheLaGlnGluSer	---	130
Db	1855	ACAGCAGCACACAAGATAAAAAGTTACGCAAGATGAATGTGCAGCGCTTGTGACTAAA	---	1914
Qy	131	MetGlyLeasnProAsnAspTyrIleProGluTyr	---	144
Db	1915	ATTAAACAATGATAAAATAATGCAATCGCAGAAATTAATAACAACAACAACAGCAAGT	---	1971
Qy	145	-----GluGlnProAsn-----	---	156
Db	1975	GTTTACAACGTAAAAAGATAACGTTTCGCAGTGTGTAGATCAAGATGTGATTACACCAACA	---	2034
Qy	157	LeuGluProGlu---LysProGlyLeuIleLysArgLeuTyrAlaArg	---	171
Db	2035	GTTTAAACCTCAAGCGAAACAAGATATTATCCAAAGCAGTTTACAACCTCGTAAACAACAAT	---	2094
Qy	172	-----LeuPheAsnAspGlyValAsnLys	---	179
Db	2095	AAAAAGTCAAAATGCATCTTACAAGATGAAAGATGTGCAAAATGATAAATTTGGTAAA	---	2154
Qy	180	ValProArgLeuLysAlaLysPheTyrGlnSerGlnSerGly	---	194
Db	2155	ATTGAACAAGGCAATTAAAGATATTGATGCAGCAACAACAATGCACAGTAGAAGCC	---	2214
Qy	195	---GluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLys	---	213
Db	2215	ATTAAACAAGGCGAATCAATGATATTAATCAA---ACTGCACCTGCTTACAACAGCTAAA	---	2271
Qy	214	---AlaAlaLeuGlu	---	217
Db	2272	GCAGCAGCTCTTGAAGAATTGACGAAGTTGTTCAGCACAATAATGTATCAAGCACTTTA	---	2331
Qy	218	-----AspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArg	---	235
Db	2332	AATCTCGATACACAATAACGAAGTAGCGAA-----GCTATTGAACGTATTAAAT	---	2382
Qy	236	GlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIle	---	253
Db	2383	GCAGCTAAAGTTTCTGGTGTAAAGCAATTGAAGCGACAAGCACTGCACAAGATTAGAA	---	2442
Qy	254	IleIleArgAsn	---	257
Db	2443	AGAGTTAAAAACGAGAAATCTCAAAATTTGAAATATTTACTGACTTACTCTCAACAAA	---	2502
Qy	257	-----	---	257
Db	2503	ATGGATGCCTATATCAAGCTTAAACAAGCAGCAACAGCTAGAAAACCTCAAAATGCTACA	---	2562
Qy	258	-----SerIleGlyGluValAspValIleIleHisAspLeuGly	---	270
Db	2563	GTTTCAAAATGCAACTAATGAAGAAGTAGCAGAAAGCTGTATGCAGAGTAGAAGCAGCTCAA	---	2622
Qy	271	GluProValTyrIleAspTyrArgAlaValGluAlaValArgGlyGluGlyAlaAspAspLys	---	290
Db	2623	AAGCAAGGTTTACATGACATCCCAAGTTGTAAATCAAAACAGGAAGTTGCTGTATACAAA	---	2682
Qy	291	-----AlaPheThrThrValAlaAspGluValProLeuLeu	---	302
Db	2683	TCAAAAGTATTAGATAAAATCAATGCAATTCACACACAGCAAAAGTTAAACCT	---	2736
Qy	303	IleGlyAspValPheHisGlyLysTyrGluThrLysAsnLeuIleGluAsnAla	---	322

Db	2737	CGAGCTGATACGGAAGTAGAAAACCGCATATAATACACGTAACCAAGAAATTCAAAATAGC	2796
Qy	323	SerAlaGlu-----HisGlyTyrPheAspGlyArgTrpLeu	334
Db	2797	AATGCTTCAACTACAGAAGAAAAACAAGCTGCATATACAGAATTAGATACTAAAAAGCAA	2856
Qy	335	AspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAsp	354
Db	2857	GAAGCAAGAAACAATCTTGATGCTGCAAAATACAAACAGTGTATGTAACAACAGCTAAAGAC	2916
Qy	355	ThrGly-----ThrGlnTyrArgPheAsp-----	362
Db	2917	AATGGTATTGCTGCAATTATCAAGTACAAACGGCGCAACAACCTAAGAAAATCGGATGCTAAA	2976
Qy	363	--GluValValPhePheThrIleAspProLysThr-----AsnGlnLeu	376
Db	2977	CGCGAATTCCTCAAAAGCAAGTGAACGTAAACTGCAATTGAAGCAATGAATCATTCG	3036
Qy	377	ThrThrAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrVal	396
Db	3037	ACTACTGAAGAACAA-----GCTGCAAAAGATAAAGTTGATCAACAGCAGTAGTTACT	3090
Qy	397	AsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAla	416
Db	3091	GCAACCGCTGATATAGATAAT-----GCTGACGCAAAATCTGATGATAGATAAT	3138
Qy	417	ThrArgTyrPheAsnMetValAsnThrGluIleValPheProGlu-----	431
Db	3139	GCAAAAACCTACTAATGAGCTACAATCGCGCCATTACCTGATGCAAAATGTTAAACCA	3198
Qy	432	-----ArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSer	445
Db	3199	ACAGCGAAACAAGCAATTGCTGATAAAGTACAAGCGCAAGAAACAGCAATTTGATGCTAAT	3258
Qy	446	SerSerSerArgThrGlu-----ProAlaGlnValAspGluSer	458
Db	3259	AACGGTGCACACACAGCAAGAAAGACGCTGCGAACCAACAAAGTTCAAACTGAAAAAAACA	3318
Qy	459	ThrLeuGluProValIleGlu-----ThrValGluLeuThrAspGlyIleLeu	474
Db	3319	ACAGCTGATACACAAATTGATGCTGCACATACAAATGACGAAGTTGAACGGCTAAAAAT	3378
Qy	475	MetaSpIleSerProIleGluPhe-----SerAlaSer	485
Db	3379	GCAGAAATTCGTAATAATGAAGCAATTCAGCCAGCAACAACAACTAAAGATAATGCGAAA	3438
Qy	486	AsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMet	505
Db	3439	CAAGCAATTGCTACGAAACGAATGAACGCTAAACAGCAATGCTGCTCAACACGCAAGACATT	3498
Qy	506	ProAspAspArgValLeuAlaIleAsnHisAspGlyValAsnArgSerIleLeuGly	525
Db	3499	ACTGCTCAAGAAATTCGACGGCGCAAT-----CGC	3528
Qy	526	ArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsn	545
Db	3529	AATGTAGATAATGCTGTGACACAAAGCAAAATAACAACATTGAAGCTGCTAATAGTCAAAAT	3588
Qy	546	GluVal-----IleAspLeuProGluArgThrAlaLeuAla	557
Db	3589	GATGTACACCAACGCAAAACAACCTGGTGAAGCTAGTATTGATCAAGTAAACACCACTACAGTT	3648
Qy	558	AsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheVal	577
Db	3649	AATAAAACACACACAGCTTACAGCGCGCAAAAATAATATT-----ACAGCTGCA	3699
Qy	578	AlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrPglySerAspThrGlyThr	597
Db	3700	ACTGATGATAATGGTGTAGTACAGCGCAAGATGCGAGGTAAAAATTCGATTCAAAGTACA	3759
Qy	598	ArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAla-----	615

```
Db 3760 CAACGACGACGCGGTAAATCAATCAATGCTAAATAAT---GATGTTGATCAAGCTGTGACA 3816
Qy 616 -----GlyAlaGluLeuArgLeuSerGluAspLys 625
Db 3817 ACTCAAAATCAAGCAATGTATATACAACTGGTGCT-----ACAACGTGAAGAAAA 3867
Qy 626 LysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArg 645
Db 3868 AATGCACCAAGATTTAGTTTAAAGCTTAAAGAAAGCGGTATCAAGATATCTTTAAAT 3927
Qy 646 AlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSer 665
Db 3928 GCA-----CAACCACTAACGAT-----3945
Qy 666 ThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArg 685
Db 3945 -----3945
Qy 686 ThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGlu-----703
Db 3946 -----GTTACGCCAAATTAAGATGTTGCGAAAGATGAATTAGCAACAAAGCA 3984
Qy 704 -----ThrTrpGlnAspLeuProValAspPheValAsnGlyLysPro 717
Db 3985 GGTATTACTGCAGATACAACTAATTAAGATGTTGCGAAAGATGAATTAGCAACAAAGCA 4044
Qy 718 ---SerGlnGluAlaLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeu 736
Db 4045 AGCGAACAAAGACATTATTGCCAAACTGCGAGTGCAGTACTACTGAGAGAAAAGAACAA 4104
Qy 737 ValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeu 756
Db 4105 GCAAAATCAACAA-----4116
Qy 757 ValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGly 776
Db 4117 ---GTAGATGCGAATTAACCGAGGTAAATCAAAATATTGAAATATGCACAGTCAATCGAT 4173
Qy 777 AspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyr 796
Db 4174 GATGTAAACACTGCAAAAGATAATGCAATTCACGAATTTGACCAATTCAGCATCAACA 4233
Qy 797 IleTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPheAlaGlyGlyAsp 816
Db 4234 GATGTTAAACGAAT-----GCAAGACGGAATTTGCTAACTGAAATGCAA 4278
Qy 817 GlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeu 836
Db 4279 AATAAAATAACTGAAATACTTAAATAATACAGACTACTAATGAAGAAAAGGT-----4332
Qy 837 ThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAsp 856
Db 4333 -----AACGAT 4338
Qy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAsp 876
Db 4339 ATT-----GCACGCTTAGAGCTGCATATGAAGAGGTTTAAATAAT---4380
Qy 877 ThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerProValGlyGlnValArgVal 896
Db 4381 -----ATTATGCGACCAACTACTACAGCTGATGTAACACT 4416
Qy 897 AspValAlaThrGlyValLysGlu-----GluGlyAsnProIleLys 910
Db 4417 GCTAAGATACAGCAGTACAAAAGATTCAACAACCTTCATGCAAAATCCTGTTAAG 4470
RESULT 5
US-09-815-242-8291
; Sequence 8291, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
```

```
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows version 4.0
; SEQ ID NO 8291
; LENGTH: 7107
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(7107)
US-09-815-242-8291

Alignment Scores:
Pred. No.: 0.00294
Score: 137.50
Percent Similarity: 29.88%
Best Local Similarity: 17.62%
Query Match: 2.91%
DB: 10
Gaps: 47

US-09-914-168-2 (1-919) x US-09-815-242-8291 (1-7107)
Qy 2 SerLysProValLeuPheAlaAsnArgSerPheMetProValAlaLeuAlaAlaTyrLeu 21
Db 1411 ACTAAGAAGTTTATCCAGATAAATCT-----TTA 1443
Qy 22 ProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIle-----38
Db 1444 AAATATCATATAAAGTTAAATGTCGAATATCGATACACCTAAAAATATTGATTTTAAAT 1503
Qy 39 -----IleAsnHisValProAlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyAsn 56
Db 1504 GAAAATTAACATATCGTACTGCTTACATATGTAATTAATAATATGCGCAA-----1554
Qy 57 ProProValLeuLeuThr-----ProGluGlnIleGlnAlaArgLeuAsnAlaAlaGly 74
Db 1555 CCAGAAGTTACACTAACATGCGAGATCCATTTTCAGTAGCAGTTGAAATGAACAAGATGCG 1614
Qy 75 LeuAsnAlaLysProGlnSerGlnAla-----83
Db 1615 TTGCAACAACAAGTAAACTCAAGTTGATAAATAGTCATTACACAGCATCAATTTGCA 1674
Qy 84 -----Leu 84
Db 1675 GAATACAATAAATTAACAACAAGCAGACACAATATTTTAAATGAAGATGCGAATCATGTT 1734
Qy 85 AspValValAsnPheAspGlnSerProIle-----95
Db 1735 GAAACTGCAAAATCGCATCTCAAGCGGCTATTGATGTTTAGTAATAAATACAAAGCT 1794
```

Qy	335	AspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAsp	354
Db	2857	GAAGCAAGAAACAATCTTGATGTCGCAAAATACAAACAGAGTATGATTAACACACAGCTAAAGAC	2916
Qy	355	ThrGly-----ThrGlnTyrArgPheAsp-----	362
Db	2917	ANTGGTATTCTGCTCAATTTAATCAAGCGGCAACAACACTAGAAATTCGGATGCTAAA	2976
Qy	363	--GluValValPhePheThrIleAspProLysThr-----AsnGlnLeu	376
Db	2977	GCAGAAATCGCTCAAAAGCAAGTGAACGTAAAACTGCAATTGAAGCAATGAATGATTCG	3036
Qy	377	ThrThrAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrVal	396
Db	3037	ACTACTGAAGAACACAA-----GCTGCAAAAGATAAAGTTGATCAAGCAGTATGTTACT	3090
Qy	397	AsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAla	416
Db	3091	GCAAACCGCTGATATAGATAT-----GCTGCAGCAAAATACTGATGTAGATAAT	3138
Qy	417	ThrArgTyrPheAsnMetValAsnThrGluIleValPheProGlu-----	431
Db	3139	GCAAAACTACTAATGAAGCTACAATCGCAGCCATTACACCTGATGCAAAATGTTAAACCA	3198
Qy	432	-----ArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSer	445
Db	3199	ACAGCGAAACAACAATTGCTGATAAAGTACAAAGCGCAAGAAACAGCAATTTGATGCTAAT	3258
Qy	446	SerSerSerArgThrGlu-----ProAlaGlnValAspGluSer	458
Db	3259	ACGGTGCACACACAGACAGAAAGACGCTGCGAACAACAAGTTCAACTGAAAAAACA	3318
Qy	459	ThrLeuGluProValIleGlu-----ThrValGluLeuThrAspGlyIleLeu	474
Db	3319	ACAGCTGATACAGCAATTGATGCTGCACATACAAATGCAGAAAGTTGAAGCGCTAAAAAT	3378
Qy	475	MetAspIleSerProIleGluPhe-----SerAlaSer	485
Db	3379	GCAGAAATGCTAAATTTGAAGCAATTCAGCGACGACCAACAACATAAGATAATGCGAAA	3438
Qy	486	AsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMet	505
Db	3439	CAAGCAATTGCTACGAAACGGATGACGTTAAACAGCAATCGCTCAACCCACAGACATT	3498
Qy	506	ProAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGly	525
Db	3499	ACTGCTCAAGCAAAATTCAGCGGCAAT-----GCG	3528
Qy	526	ArgIleSerAspIleValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsn	545
Db	3529	AATGTAGATTAATGCTGTGACACAAAGCAAAATACACATTTGAAGCTGCTTAATAGTCAAAAT	3588
Qy	546	GluVal-----IleAspLeuProGluArgThrAlaLeuAla	557
Db	3589	GATGTAGACCAACGGAACACACTGGTGAAGCTAGTATTGATCAAGTAAACCAACACAGTT	3648
Qy	558	AsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheVal	577
Db	3649	AATAAAAAACACACAGCAGTTCACAGCGGCAAAATATATT-----ACAGCTGCA	3699
Qy	578	AlaSerAspLysProArgAspGlyIleGlyLeuGlyTyrPglySerAspThrGlyThr	597
Db	3700	ACTGATGATTAATGGTGTAGATACAGCGAAGATGCAGGTAAAAATTCGATTCAAAGTACA	3759
Qy	598	ArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAla-----	615
Db	3760	CAACGACCAACACGGTTAAATCAAAATGCTAAAAAT---GATGTTGATCAAGCTGCACA	3816
Qy	616	-----GlyAlaGluLeuArgLeuSerGluAspLys	625
Db	3817	ACTCAAAATCAACCAATTGATTAATCAACTGGTGT-----ACAACCTGAAGAAAAA	3867
Qy	626	LysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArg	645

Qy	461	uProValleleGluThrValGluLeuThrAspGlylleLeuMetAspIleSerProIleGl	481
Db	24044	AGGACAGATTGCTGCTGCCAGACAGATTGCTGCGAGGCAAT	24080
Qy	481	uPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaLysAlaArgHi	501
Db	24081	---GCCACGGAAGGATGACAGAAGCGGTAAACAGCTTCTGTGCTCTGCACGACGC	24134
Qy	501	sLeuTyArgMetProAspAspArgValLeuAlaIleAsnHisAspAspGlylValAsnAr	521
Db	24135	-ATCAGCGACCTGCGCGGAAAAAACTCACGGCA-----CATGAA-----AA	24175
Qy	521	gSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAs	541
Db	24176	GAGTGTCTGCCCGTAAAGATGAACATGATTTCAGCGCTGACG-----CTGTGGA	24226
Qy	541	pGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysTh	561
Db	24227	TGTAACACAGCAGAG-----CTTCAGAACACACGCGCACTCAACGACTGAAG--	24275
Qy	561	rProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLy	581
Db	24276	-----AAAAAACAAATTCAGCTGACCACTGCTGAGTCACTGCGTGAAGAAGA	24316
Qy	581	sProArgAspGlyGln-----IleGlyLeuGlyTrpGlySe	593
Db	24317	CGCGCTGACGTACAGCAACATGACCTCGATATCCGCCACGCTGGGTATG-----	24365
Qy	593	rAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTy	613
Db	24366	-----GGTGATCAGCAGCGACGCATATCAGGTACACTGAGTCTGCCAGAAATA	24418
Qy	613	rGlnAlaGlyAlaGlu---LeuArgLeuSerGluAspLysLysGlyValLysLeuTy--	631
Db	24419	CCAGCAACAGCTGCAGCAGTTGAGCGGGATGAGCAGAGAAGGGACA-----TATAA	24472
Qy	632	-----AlaThrLysProLeuSerHisProLeuAsnAspGlnLeuAr	645
Db	24473	CACGGATGACTACAGAAGCGCAGCGCTGACGGAGAGCCGTGAACGCCAACCTGAA	24532
Qy	645	gAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSe	665
Db	24533	TGAGAATCGCGTTACTGGCAACAGCTT-----	24560
Qy	665	rThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnAr	685
Db	24561	-----GAAGTTGTACAG---GGTAACTCG---AA	24583
Qy	685	gThrTySerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTr	705
Db	24584	AAACGGAGTCTCGCT-----GCATT	24604
Qy	705	pGlnAspLeuValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGl	725
Db	24605	TCAGATTTTACCGTGGATGCAGATATATACGGCAGAAACACAGCAA-----	24650
Qy	725	yValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGl	745
Db	24651	---CAGTGTCCTCGTCAGCGCTTCAGCAACATCGGAATGGCCCTGCGCACTTTGT	24706
Qy	745	nArgTySerLeuGluValGlySer-----SerGlyLeuValSerAspAlaAsnMetAl	763
Db	24707	TACGGCAAACTCAATTTCAAATCCTTCACCTCTCTGTGTGTGATATGGCGAAAT	24766
Qy	763	alleAlaArgAla-----GlyIleSerGlyValTyrSerPheGl	776
Db	24767	CCTGGCGCAGCAACCATGATGAATCGATAAAGGATGGCAGTGTACTGGGATTGGA	24826
Qy	776	yAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTy	796
Db	24827	TCTCAGCAGCCTTCCCTGAT-----GCCAATGGGGGATTTATCAGTCTGC	24874
Qy	796	rIleTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPheAla	813

```
Db 24875 TCATTGAGTCGTTACAGTGGCACCGTGGTTAAACGTCGCGAGTTTGTCTTTTGCAAA 24934
Qy 814 -----GlyGlyAspGlnSerIle-----ArGlyTyrAl 823
Db 24935 AGCGCGGTGTGATGGGGAAGCGGACCTGAACCATTCGCCATTCGCGTGGTGGTGC 24994
Qy 823 ahisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlnValLeuAl 843
Db 24995 TCACGCTAAGCTGGGGTGTGGCGGAT-----ATTGGGGTTCAGGTATGGC 25042
Qy 843 aValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGl 863
Db 25043 GATGTTTCCCGCGAGTAAACATCGAGATCAATAACGAT-----GCCACGAACGG 25093
Qy 863 yAspIleGlyAsnAla-----TyrAspLysGly 872
Db 25094 GCAGATAGTCCGGCTGCCTGAAGCGGTTTATGACCTCGGG 25136
```

RESULT 7

```
US-09-815-242-7474
; Sequence 7474, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7474
; LENGTH: 8673
; TYPE: DNA
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(8673)
; US-09-815-242-7474
```

```
Alignment Scores:
Pred. No.: 0.0057 Length: 8673
Score: 136.00 Matches: 198
Percent Similarity: 33.56% Conservative: 145
Best Local Similarity: 19.37% Mismatches: 361
Query Match: 2.88% Indels: 318
DB: 10 Gaps: 52
```

US-09-914-168-2 (1-919) x US-09-815-242-7474 (1-8673)

Qy 55 GlyAsnProValLeuLeuThrProGluGlnIleGlnAlaArgLeuAsnAlaGly 74

```
Db 5023 GGACGCCAGAAATCATTTGCGCAATGAAAA----- 5055
Qy 75 LeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspGlnSerPro 94
Db 5056 -----AGGATGTTGCAAGAGCGGTGATGCTGCTTTTATACAGCCGCTAGCACC 5106
Qy 95 IleSerArgIleGlyGluGlnSerProProLeuGlyLeuAspMetSerValIleGluGlu 114
Db 5107 AATGCGGTAAAGGGGTAAACAACGCGCTTTA----- 5139
Qy 115 ThrThrProLeuSerLeuGluGluLeuPheAlaGlnGlnSerThrGluMetGlyIleAsn 134
Db 5140 -----AAATCGCTCAGCGAGATCATTAAGGCAAGCAAGGCGCTTTCAGCGCAAAAC 5190
Qy 135 ProAsnAspTyrIleProGluTyrGlnGlyGluGlnProAsnSerGluValValPro 154
Db 5191 CTTTATAGTAAGCGCGGTGATTTTTCAGGC-----AGAAGCGTGTGTCGTGGGG 5241
Qy 155 ProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsn 174
Db 5242 CCTAACCTTAAATGATGATCGGATTCGCTTAAAAACATGCGGTAGAACTCTTCAAA 5301
Qy 175 AspGlyValAsnLysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGly 194
Db 5302 -----CCGCATCTGTTATCCACGCTT-----GAAGAGACAGCG 5334
Qy 195 GluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsn-----IleLys 213
Db 5335 TATGCCACCCACGCTCAACAGGCTAAACAGCATGATTTGACAAAGAGCAATGAAGTGTGG 5394
Qy 214 AlaAlaLeuGluAspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArg 233
Db 5395 GAGTGCTTGCAGAAATCAGAGGGGTATCCGGTGCTACTCAACCGGCC---CCTACC 5451
Qy 234 LeuArgGlnThrAla-----LeuValAlaAlaArgAlaValGly 246
Db 5452 TTGCACAAGCAATCCATTCACGCTTCCATCCAAAGCTCATTCAGCGCAAGCGATCCAG 5511
Qy 247 TyrTyr-----AspIleAsp----- 251
Db 5512 TTGCACCCGTTAGTGCTCAGCGTTTAAAGCGTGTGTTGACGGGGATCAATGCGGTG 5571
Qy 252 ---LeuSerIleIleArgAsnSerIleGlyGlu-----Val 262
Db 5572 CATGTGCTTTAAGCCAGGAAGCGATCGCTGAATGCAAGGTCTCATGCTAAGCTCTATG 5631
Qy 263 AspValIleIleHisAspLeuGlyGluProValTyrIleAsp----- 276
Db 5632 AATATCCTTTTACCCGCTAGCGGTAAAGCGGTAGCCATTCCTAGCCAGGATATGTTTTA 5691
Qy 277 -----TyrArgAlaValArgGlyGlyGluGlyAlaAspAspLysAlaPheThr 293
Db 5692 GGGCTTTATTATCTTTTACAAAAGAGTGGGTCAGGCGGAGCATTAAGCTTTCTCT 5751
Qy 294 ThrValAlaAspGluValProLeuLeu-----IleGlyAspValPheHisGlyLysTyr 312
Db 5752 AGCGTGAATGAATCATCACCGCTATTGACACGAAAGAAATTAGACATCCACCAAGATT 5811
Qy 313 GluThr-----LysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAsp 330
Db 5812 AGGGTTTTGGATCAAGGGAATATTATCGCTACGACGCGG----- 5850
Qy 331 GlyArgTrpLeuAspArgSerValAspValIleLeuProAsp----- 344
Db 5851 GGGCGCATGATCATCAAGTCC-----ATTTCCTGATTTTATCCCTACCGATTG 5901
Qy 345 -----AsnThrAlaAspValSerLeuIleTyrAsp----- 354
Db 5902 TGGACACACCCATGAAGAAAAACATATTGCGTGTGTTGGATTATGTCATAAAGTG 5961
Qy 354 ----- 354
```


Db 5962 GCGGATATCGCATACCGCAACCTTTTGGATCATCTTGAACCGCTTGGCTTATAGGAT 6021
Qy 355 -----ThrGlyThrGlnTyArgPheAspGluValValPhePheThrIleAspPro 371
Db 6022 CGCACTAAGCGTGGTATTTCTATCTCTATGGAAGATATT-----ATCAGCGCA 6069
Qy 372 LysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArg-----GluLeu 389
Db 6070 AAAGACAAGCAAAATGCTGGAAAAAGCCAAAGTAGAAGTTAAAAAATCCAACAACAA 6129
Qy 390 LeuGluGlnLeuThrValAsnMetGlyGluAlaTyAsnLeuGlnAlaValArgAla 409
Db 6130 TACGATCAGGGTGTCTCCTACCTGACAA--GAGCGTTAC----- 6165
Qy 410 LeuSerAsnAspLeuIleAlaThrArgTyPheAsnMetValAsnThrGluIleValPhe 429
Db 6166 -----AATAAAATCATTGACACT-----TGGACTGAAGTCAATGACAGAATGAGTAAG 6213
Qy 430 ProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGln----- 444
Db 6214 GAAATGATGAGCGGTATCGCAAAAGATAAAGAGGGCTTTAAACTCTATTATATATGATGCG 6273
Qy 445 SerSerSerArgThrGluProAlaGlnVal----- 455
Db 6274 GATAGTGGCGCAAGGGGAGCGCGCGCAATCCGTCAGCTTTTCAGCGATGAGGGGGCTT 6333
Qy 456 -----AspGluSerThrLeuGlu---ProValIleGluThrValGluLeuThr 470
Db 6334 ATGACAAAGCGGATGGCAGTATCATTTAAACGCCCATATTATTCTTAACTTTAAAGAGGG 6393
Qy 471 AspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAsp 490
Db 6394 TTGAATGCTTTAGAAATCTTTAACTCCACGATGCGGTAGAAAGGGCTTAGCGGATACA 6453
Qy 491 LysLeuAsnLeuValAlaAla-----LysAlaArgHisLeuTyArgMetProAspAsp 508
Db 6454 CGCTAAACAGCAATGCGAGGTATTTTGACAGAAACTCATTTGACGTT---TCGCAA 6510
Qy 509 ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSer 528
Db 6511 AATGTCAAAGTGGTCTCTGATGATTCGCGCACCGCATGAAGGATT-----GAAATCACG 6564
Qy 529 AspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGlu----- 544
Db 6565 GAT---ATTGGCGTGGGAGTGAGTGATGAACCTTTAGAAAGCGGTATTTTCGGGCGC 6621
Qy 545 -----AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThr 561
Db 6622 GTTTTATTAGAGATGTGATCGATCCCATTTACCAATGGAATCTTGCTCTATCGGACACT 6681
Qy 562 ProAlaAspValTyArgLysLysVal-----ProLeu 573
Db 6682 TTGATTGACGAAGAGGGTGTAAAGAGTGGTTGAAGCTGGGATTAATCCATTACGATC 6741
Qy 574 TyValPheValAlaSerAspLysProArgAsp-----GlyGlnIleGly 588
Db 6742 CGCACCCAGTAAGTGTAAAGCGCCAAAGGCGGTGTCGCCGAATGCTATGGCTTGAAT 6801
Qy 589 LeuGlyTyArgGly----- 592
Db 6802 TTAGCGAAGCAAGATGAGCTATCCGGGTGAAGCGGTGGGCGGTAGCCGCGCAATCT 6861
Qy 593 ---SerAspThrGlyThrArgLeuValThrLysPheGluHis----- 605
Db 6862 ATCGGGGAGCGCTGGAACCGACGCTCACTTTAAGGACTTTTCCATGTGGCGGACACGAGC 6921
Qy 606 -----AsnLeuIleAsnArgAspGlyTyArgGlnAlaGlyAlaGlu 618
Db 6922 AGGAGTCAGGATGACGCGGAATATCGTGGCGAGCAAGAGGTTTTGTGCGTTCTTACAAC 6981
Qy 619 LeuArgLeuSerGluAspLysLysGlyValLysLeuTyAlaThrLys----- 634
Db 6982 CTTAGGACTTACAGAAATAAAGGGTAAACATTTATCGCCAAACCGCGCTTCT 7041

Qy 635 -----ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeu 648
Db 7042 ATTTTACTGCTAGACCTTAAGATTAAAGCGCTTTTGATGGGAATTACGCATTCAAACG 7101
Qy 649 GlyTyArgGlnGlnGluValPheClyHisSerThrAsnGlyPheAspLeuSerThrArgThr 668
Db 7102 GTCTATGAAGAGTGTGTTGAGC----- 7125
Qy 669 LeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyPheAsnArgThrTySer 688
Db 7126 -----GTGAAAAATGGGATCAAGACCTTAATTCGTT 7158
Qy 689 LeuArgTyArgLeuAspLysLeuLysThrGlnAlaProGluThrTrpGlnAspLeu 708
Db 7159 TTAAGA---AGAAGCATATTGTCAAG-----CCAAGCGAA----- 7191
Qy 709 ProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeu-----AlaGly 725
Db 7192 ---TTAGCGCGCTTGGCGGTAAGATTGAGGGGAAAGTGTATTTGCCCTTATGCTAGCGG 7248
Qy 726 ValAlaValHisLys-----ThrValAlaAspAsnLeuValAsnProMetArgGlyTy 743
Db 7249 CATAGGTGCATAAGGGGGAAGTATCGCTGATATTATC-----CAAGAGGCGTGG 7299
Qy 744 ArgGlnArgTySerSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
Db 7300 AATGTGCCTAATCGCATCCCTTATCGAGCGAATTCGTGTTAAAGAT---AATGACCC 7356
Qy 764 IleAlaArg-----AlaGlyIleSerGlyValTySerPheGlyAspAsnAlaTy 780
Db 7357 ATCGCGCAAGATGTGTATGCCAAAGAAAAGGTGTGATCAAAATAC-----TAT 7404
Qy 781 GlySerAsnArgAlaHis---GlnMetThrGlyGlyIleGlnAlaGlyTyTrpSer 799
Db 7405 GTTTTAGAGGCCAATCATTTAGAGCGCACCATGAGTCAAAAAGGGCGATATTGTGAGT 7464
Qy 800 AsnAspPheAsnHisValProTyArgLeuArgPhePheAla-----Gly 814
Db 7465 GAAAAA-----GGCTTTTTCAGTGTAGTGCATGATAAT 7500
Qy 815 GlyAspGlnSerIleArgGlyTy---AlaHisAspSerLeuSerProIleSerAspLys 833
Db 7501 GGTAGGAAGCGCGTCGCATTTATATCGCTAGGGTTCTGAGATCTTGATTGATGATAAT 7560
Qy 834 GlyTyLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGlu----- 848
Db 7561 AGTGAAGTGAGCGCTAATAGTGTAAATTTCTAAACCCAGCACTTAACACTTTCAAACGATT 7620
Qy 849 -----TyAsnTyArgGluPheMetLysAspLeuArgLeuAlaValPheGly 863
Db 7621 GCCACATGGGATCCCTTACAACACCCCTATCATTTGCGGACTTT----- 7662
Qy 864 AspIleGlyAsnAlaTyArgAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyVal 883
Db 7663 -----AAAGGTAAGGTGAATTTTGTGGATGTTATCGCAGGGGT 7701
Qy 884 GlyVal 885
Db 7702 ACGGTC 7707

RESULT 8

US-10-108-605-236
; Sequence 236, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOCASTER THAT ENCOD
; TITLE OF INVENTION: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; FILE REFERENCE: 31133B

; CURRENT APPLICATION NUMBER: US/10/108.605
 ; CURRENT FILING DATE: 2002-03-27
 ; PRIOR APPLICATION NUMBER: US 07/61,142
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/176,418
 ; PRIOR FILING DATE: 2000-01-14
 ; NUMBER OF SEQ ID NOS: 361
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 236
 ; LENGTH: 12404
 ; TYPE: DNA
 ; ORGANISM: Drosophila melanogaster
 US-10-108-605-236

Alignment Scores:

Pred. No.: 0.0094 Length: 12404
 Score: 136.00 Matches: 132
 Percent Similarity: 31.54% Conservative: 103
 Best Local Similarity: 17.72% Mismatches: 253
 Query Match: 2.88% Indels: 257
 DB: 9 Gaps: 32

US-09-914-168-2 (1-919) x US-10-108-605-236 (1-12404)

Qy 14 ProValAlaLeuAlaAlaValLeuProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsn 33
 Db 1966 CCGACTATAATGCTGCTCTCTGAGTGCAGTCCCTCCAAAATATCGGAGTCGGCC 2025
 Qy 34 AsnProAlaAsnIleAlaAsnHisValProAlaHisAspThrAlaIleAsnGlnAlaLys 53
 Db 2026 CAACCGGGGGAATCGTGGCCAGAATTCGGTGCATGAC----- 2064
 Qy 54 AlaGlyAsnProValLeuLeuThrProGluGlnIleGlnAlaAargLeuAsnAlaAla 73
 Db 2065 -----CCCGATTTCMAAACCGAGTATGCCAATGTGAATGACGCTCAATGAGGA 2115
 Qy 74 -----GlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPhe 89
 Db 2116 GATGTCCTACTTCGGCTTGACCAACCGAGTAACAGTATTATTGTTATAGTTCACCTG 2175
 Qy 90 Asp--AspGlnSerProIleSerArgIleGlyGlnSerProProLeuGlyLeuAsp 108
 Db 2176 CCGCTGCACGGGAGATGTTCA-----AATTATACA 2208
 Qy 109 MetSerValIle-----GluGluThrThrProLeuSerLeuGluGluLeuPheAla 125
 Db 2209 CTGAGTCTGTAGCCACCGATAAGGAACTCCGCCCTGCACCCCTCCAATCGATCTC 2268
 Qy 126 GlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGlu 145
 Db 2269 CTAAGGATTACCGATGTCATGCAATCCA-----CCGGAGTTCGAGCAGGAC 2316
 Qy 146 GlnProAsnSerGluValValProProThrLeuGluProGluLysProGlyLeu--- 164
 Db 2317 CTCTACCAACGCCACGTC-----ATGAGGTGGCCGCGGTACATCG 2361
 Qy 165 IleLysArgLeuTyrAlaAargLeuPheAsnAspGlyValAsn----- 178
 Db 2362 GTACTTCAAGTTCTAGCCACCGATCGAGATGAGGCGCTAAACTCGGCTTGACATACTCG 2421
 Qy 179 -----LysValProArgLeuLysAlaLysPheTyrGln---SerSerGlnSerGlyGlu 195
 Db 2422 CTGGCCGAACCGCCGAACCCATGCCAGTGGTTTCAGATTGACCCACAAACGGGCGCTG 2481
 Qy 196 ThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAla 215
 Db 2482 ATCACCACCGGATCGCATATCGATTGCGAAACGAGCCGGTCGCCGAGCTAACGGTTGTC 2541
 Qy 216 LeuGluAspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArg 235
 Db 2542 GCTAGGGAC-----GGAGGTGTCGCCGCCCTTCC 2571
 Qy 236 GlnThrAlaLeuValAlaAlaAargAlaValGlyTyrAspIleAspLeuSerIleIle 255

Db 2572 TCCACAGCCACGGTCTT----- 2589
 Qy 256 ArgAsnSerIleGlyGluValAspValIleIleHisAspLeu-----GlyGluProVal 273
 Db 2590 -----GTTACCATTCCAGCTGTAACACGACACGACCAACCGCAT 2625
 Qy 274 TyrIleAspTyrArgAlaValGluValArgGlyGluAlaAspAspLysAlaPhe--- 292
 Db 2626 TTC-----GACCAGAGCTTCTAC 2643
 Qy 293 ---ThrThrValAlaAspGluValProLeu----- 301
 Db 2644 AACGTATCCGCTCGCTGAAACCAACCGCTGGCAGGTGCATCTCAAGGTCTCTGCAAGT 2703
 Qy 302 -----LeuIleGlyAspValPheHisHis 309
 Db 2704 GATCCAGATTGGGAGTAATGCCATGGTGAACCTACACCATCGCGAGGCTTCAAGCAC 2763
 Qy 310 ---GlyLysTyrGluThrLysAsnLeuIleGluAsnAlaSerAlaGlu----- 325
 Db 2764 CTCACCGAATTCGAG-----GTGGCTCCGCTTCTGCGAGATTTCGATA 2808
 Qy 326 HisGlyTyrPheAspGlyArgTrpLeuAspArgSerValAspValIleLeuProAspAsn 345
 Db 2809 GCTGGCGAATTCGAC----- 2823
 Qy 346 ThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValVal 365
 Db 2824 -----TTCGAGCGGAGGAGTAGCTATGAATTTCC- 2853
 Qy 366 PhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProVal 385
 Db 2854 -----GTATGGCAACCATCGTGGGGTCTTGAGCACT 2886
 Qy 386 LysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetClyGluAlaTyrAsnLeuGln 405
 Db 2887 ACTGCCATGATTAATGCAATTGACGGAGCTGAAC----- 2922
 Qy 406 AlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThr 425
 Db 2923 -----GACATCGT 2931
 Qy 426 GluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSer 445
 Db 2932 CCGGTTTCTATCCCGAG-----GAATATAA 2958
 Qy 446 SerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGlu 465
 Db 2959 GTATCGCTCAGGAATCCCGAAGAGCTCTCACAGCTTCGAGTACTCGGATTGCGCT 3018
 Qy 466 ThrVal-----GluLeuThrAsp 471
 Db 3019 GTGTGGCCACAGATCCGGATTACGGAACTTTGGCCAGGTCTCATATAGGATTGTGGCC 3078
 Qy 472 GlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLys 491
 Db 3079 GGTACGACGGCGGCATCTCCGGATTGATCGATCCACTGGTGAATATTCGTGTCGCT 3138
 Qy 492 LeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeu 511
 Db 3139 CCGGATATGCTATCCGTTCGTACCCCAACCCATGAC-----ATGCTA 3180
 Qy 512 AlaIleAsnHisAspAspGlyValAlaArgSerIleLeuGlyArgIleSerAspAlaVal 531
 Db 3181 AATATATCGCGCACCGATGCGCGTAAT-----CTAAGGAGCAATGCCACCGCGTG 3231
 Qy 532 SerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuPro 551
 Db 3232 GTCTCTTCAGC-----ATCATCCACGCCCATG 3258
 Qy 552 GluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysVal 571

Db 3259 CAGCGACCGCGATC-----TTCAAAGGCC 3285
Qy 572 ProLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGly-----Leu 589
Db 3286 AGATACAATTACTGTGCAAGGACACATTCGAGGGCGCACGGTTGTGGTTCTGTGATA 3345
Qy 590 GlyTyrGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsn 609
Db 3346 GCTCGGAGTGAGATGTCGCCCATCGAAGTCGGTTAGGTATTCGATATACATCGGGCGCAT 3405
Qy 610 ArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLys 629
Db 3406 CCCGATGGCTAT-----TTCAAGTAAAGAACCACTCTGGAAC 3444
Qy 630 LeuTyrAlaThrLysProLysProLysHisProLeuAsnAspGln-----LeuArg 645
Db 3445 ATACGTATAGTAAAGCCCTGGATCAAGCGAAGTCCCGAGTTCTACTCAACATCCAA 3504
Qy 646 AlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSer 665
Db 3505 GCCACCTTGGG---GAACCTCCGGTTATGGACACACAGGTTAATATCGAAGTGGAG 3561
Qy 666 -ThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnAr 685
Db 3562 GATGTCACAGATATGCACCTGAGTTCGAGGCCAGTATGGTTAGGATTTCCGTACCGGAA 3621
Qy 685 gThrTyrSerLeu 689
Db 3622 AGTCAGAGCTTG 3634

RESULT 9

US-10-114-170-243/c

; Sequence 243, Application US/10114170

; Publication No. US20030023075A1

; GENERAL INFORMATION:

; APPLICANT: Blattner, Frederick R.

; Burland, Frederick R.

; Perna, Nicole T.

; Plunkett, Guy

; Welch, Rod

; TITLE OF INVENTION: No. US20030023075A1el sequences of E. coli O157

; NUMBER OF SEQUENCES: 265

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Charles & Brady

; STREET: 1 South Pinckney Street

; CITY: Madison

; STATE: WI

; COUNTRY: US

; ZIP: 53701-2113

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Word Perfect 8.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/114,170

; FILING DATE: 01-Apr-2002

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/453,702

; FILING DATE: 03-DEC-1999

; APPLICATION NUMBER: 60/110,955

; FILING DATE: 04-DEC-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Seay, Nicholas J.

; REGISTRATION NUMBER: 27386

; REFERENCE/DOCKET NUMBER: 960296.95017

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (608) 251-5000

; TELEFAX: (608) 251-9166

; INFORMATION FOR SEQ ID NO: 243:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 5629

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 243:
US-10-114-170-243

Alignment Scores:

Pred. No.: 0.00373 Length: 5629
Score: 135.00 Matches: 205
Percent Similarity: 31.34% Conservative: 157
Best Local Similarity: 17.75% Mismatches: 365
Query Match: 2.86% Indels: 428
DB: 9 Gaps: 51

US-09-914-168-2 (1-919) x US-10-114-170-243 (1-5629)

Qy 54 AlaGlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaAargLeuAsnAlaAla 73
Db 3289 GCCCGCAATTATCCGGCTATAAGCAAGGGAGCTGACATAAGGGCACAACTTAGCCTTCAC 3230
Qy 74 GlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspAspGlnSer 93
Db 3229 AACTGGAATGCTCCACTCCAATCCCATATTTATACATTGAGGCAACCAACCAATAAGCT 3170
Qy 94 ProIleSerArgIleGlyGluGlnSerProProLeuGlyLeuAspMetSerValIleGlu 113
Db 3169 CGCGTTGCCACATTATCAGCGACAAATAATGACGCTTTACGCCGATAAAAACACATTTAAT 3110
Qy 114 GluThrThrProLeuSerLeuGluGluPheAlaGlnGluSerThrGluMetGlyIle 133
Db 3109 ACCCTCACG-----ATCAACGTCACCTGATGAGAGTGAT 3077
Qy 134 AsnPro---AsnAspTyrIleProGluTyrGlnGlyGluGlnProAsnSerGluValVal 152
Db 3076 AATCCCTCGCAAAATCATCAGGTCACCTTAAAGAAATGAAAGGAGGCGGAGTTGTC 3017
Qy 153 ValProPro----- 155
Db 3016 GAACCGCGCAGCAAAATACGGATGCATATGGTGTGCCACAATAAACATGTAAGTCAG 2957
Qy 156 -----ThrLeuGluProGluLysPro---GlyLeuIleLysArgLeuTyr 169
Db 2956 GTTCGGAAGAAATACGATTAGCGCCACGCTGCCAAATGTTTTCACACGGATAATT 2897
Qy 170 AlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyrGln 189
Db 2896 GCGAAATTCGTTAGCGAT-----TCGAGTACGCCA-----AAATTCACAA 2855
Qy 190 SerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr----- 206
Db 2854 CTGTTTCCCGATCCAGATACCATTTATTGCTGGCAACAGCAGGCGAGTACTCTGACGCC 2795
Qy 207 -----GluProTyrAlaAsnIleLys----- 213
Db 2794 ATCATCACAGACTTTTCATACACACCCGTTAAAGATGATGAAGTGAATTTTGTGCGACCT 2735
Qy 214 -----AlaAlaLeuGluAspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIle 231
Db 2734 GGTGGCTCGCAACTGGACACACACGCC-----GCCACACAGACACGCTCCGGTATTGTG 2681
Qy 232 ProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAsp 251
Db 2680 -----CGGGTGCACCTGACCGATTCAAAAGCT---GGTAGCTATTTCGCTCGAT 2636
Qy 252 LeuSerIleIleArgAsnSerIleGlyClyValAspValIleIleHisAspLeuGlyCly 271
Db 2635 GCCTCGCTT-----GAGGTGGATAAAATAATTCACCACTCGGTGTCAG 2594
Qy 272 ProValTyrIleAspTyrArgAlaValGluVal-----ArgGlyGluGlyAla 287
Db 2593 ATCACCCTGGTCCCAACACAGGAAACAACTCGGTAAATGACCTTGAATGCCGGGTGCGGCGAT 2534

QY 287 ----- 287
Db 2533 GCGATCGCTAAACAATACAAATATCGTTACCTGACTGCCAGTGTGAAGATGTTTATGGA 2474
QY 288 ----- 296
Db 2473 CACCCGTTGCCGGATGAGGATGTGAATTTACCTTGCCAGCCTCCATGACCGGGAACITC 2414
QY 297 ----- 303
Db 2413 ACGCTAAGTAGTGAACCGCCGACCGATGCAACAGGTGATGCCGTGGTCACATTCGGA 2354
QY 304 GlyAspValPheHisHisGlyLysThrGluThrLysLysAsnLeuIleGluAsnAlaSer 323
Db 2353 GGCACA-----AAAGCGGGTGAGTTTACAGTTACGGCGAGCTTGACCCAGAAATAATACC 2300
QY 324 AlaGluHisGlyTyr-----PheAspGly----- 334
Db 2299 GTTGCTTATCAGCAAGTCACTTTTATTGGGATACAAACAGTGGCGAGCTCCAGCCGCTG 2240
QY 335 AspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeu----- 351
Db 2239 ACTGCCTCATTAATTCATTTGTCGGGTAAACAGTACGGGGAGTACCCTGACGCGCAACG 2180
QY 352 IleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePhe----- 367
Db 2179 ATCCTGGACGCTTACCAAAATCCGCTTAAGACGAGTTGGTCACTTTCCACAGTAACGAT 2120
QY 367 ----- 367
Db 2119 GTCACCTCTAAGCGAAACAGAAAGTCAACCACCAATACGCTGGTCAAGCGACGGTAACAATG 2060
QY 368 ----- 370
Db 2059 ACCAGCAATATTGCCGGACAAACATAAGCTGCTGGTGAAGCGGAAAGCGCAAGCTTCCGAT 2000
QY 371 ProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLys-----ArgGluLeu 389
Db 1999 AATAAAACGTTTAGTTTATCAGTCTACCGGATCAAGTTCGGCGAAGTGAATAAGTATA 1940
QY 390 LeuGluGlnLeuThrValAsnMetGlyGluAlaTyrAsnLeuGln-----AlaValArg 408
Db 1939 ACCGAGCCGCAAAACGATACAGGTGGCGGAAACATCACGCTACGGAATACTCGTCCAG 1880
QY 409 AlaLeuSerAsnAspLeuIleAlaThrArg----- 418
Db 1879 GACCGGTTTACAATGTAAATCGCGGGTCAACGGCTCAGATTAGTGGCGACGCCAACAAT 1820
QY 419 ----- 422
Db 1819 AACATTACGATAGCGGATACGGCTTACACCGATAATAACCGTTATGCGTACGTTAACCTT 1760
QY 423 ValAsnThrGluIleValPhePrdGluArgGluGlnIleGlnAsnAspGlnValSerPhe 442
Db 1759 CTCAGCACCCAA-----CTGGGGTTTATCAGGTG-----ACGCGAAGCGTG 1718
QY 443 GluGlnSerSerSerArgThrGlu----- 451
Db 1717 GACAAATAACAGTAGTAGTAGTTGACGTGAATGTGGCAATGGCAAACTCGAGTTAACA 1658
QY 451 ----- 451
Db 1657 TCATCGAAACCCAGAAACTACGGTCCATAATAGTAGGGTATTACGCTGACCGCAACGGCG 1598
QY 452 ----- 463
Db 1597 AGAATGCGCGGGGTGAATGTATGTCAGCGCAAAATTTATCACCTTTAGCGTAACGCGCTGAA 1538
QY 464 IleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSer 483
Db 1537 GGTGCAACGCTAGCAATACAGGGGAAGTCTTACTGACCACTCA-----GTCAGGCC 1484
QY 484 AlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaAlaArgHisLeuTyr 503

Db 1483 AAAGTAGCGTGCACCTGACAAAGTGAATGTTCTATACCGTTACGGCCATAATGGGCAAA 1424
QY 504 AspMetPro-----AspAspArgValLeuAlaIleAsnHisAsp----- 516
Db 1423 GATGTTCCCGTTTCAGAGCCAGGTAAACGGTTTAAAGCAGCATGCTAAACAGCGCAT 1364
QY 517 ----- 522
Db 1363 GTTGTGACCGTCGTGGGTTCTCTGACACCATCACCAGCCGCGCATCGATACGACGACC 1304
QY 523 IleLeuGlyArgIleSerAsp-----AlaValSerAlaValAlaArgAlaIle 538
Db 1303 ATCATCTCAGAGTAGAAGATGATTACGGATTCCCGGTTGAAGGTGTCGATATTAGTCAT 1244
QY 539 LeuProAspGluSerGluAsnGluValIleAspLeuPro----- 551
Db 1243 GGCTTAGACACCAAGCAGCCGCGTAGTTAATATTCCAACCTACGGGTACCGATCAGTCC 1184
QY 552 ----- 560
Db 1183 GGCAAGTCAACGGCGACAATAACCAAGTACATTGGCAGAAACCTTAACAGTCAATGTGCAA 1124
QY 561 ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAsp 580
Db 1123 GTTCCTGGCACAGCCCAACCAATCCGCAACCATTTACATTGTTGCCGCGACGCCGATGAA 1064
QY 581 LysProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuVal 600
Db 1063 ----- 1052
QY 601 ThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArg 620
Db 1051 TTCAAAATCCGATCTTCGACTCTGAAAGGTGAC-----TACCAGCAGAGCGCAAACTTACG 995
QY 621 LeuSer-----GluAspLysLysGlyValLysLeuTyrAlaThrLysProLeu----- 636
Db 994 CTAAATTCGACAGCAAGTACGGTAACCCGATAGTAGCTCTGATCATCTGGAATTTGTC 935
QY 637 ---SerHisProLeuAsnAspGlnLeuArg---AlaThrLeuGlyTyrGlnGluVal 654
Db 934 CAGTCAGGCCCTTCGTGAACTTTCACAGTTCAGCGGATATTGATTACAGCCCAAGAAAT 875
QY 655 PheGlyHisSerThr----- 659
Db 874 TATGGCGAGTACACCGCTGACTGCTGCGGAAAGAGGAAACAGCAGACTCATTTCC 815
QY 660 ----- 677
Db 814 ATGCTGAACGGGTTTCATCAGGCAAACTTAAGCATATATCGCTGAATCTCATCAATCGATA 755
QY 678 IleGlnAsnGlyGlyTyr-----AsnArgThrTyrSerLeuArgTyrArgLeu 693
Db 754 AAAGAAATGTCGCGTTCATGCTCACTGCAACACCACTACCTCTCTCCAGCGCT----- 704
QY 694 AspLysLeuLysThrGlnAlaProProGluThrTyrGlnAspLeuProValAspPheVal 713
Db 703 ---AAATCCCGAGCGAAGGCTTTGCAGGAGCGTATTACACACTC----- 662
QY 714 AsnGlyLysProSerGlnGluAlaLeuLeuAlaValAlaValHisLysThrValAla 733
Db 661 ----- 626
QY 734 AspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySer 753
Db 625 GATTATATGTTTTCAGTTTCACAGGGTTGG-----GTGCTGCTCCAGTCT 581
QY 754 SerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyr 773
Db 580 TCGGTAAGTTTCTTCGCAAAATATC----- 554
QY 774 SerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGly----- 791

Db 553 -----GGCGATCAACGTCAGTCACAATAAGCGCTGTCCCGACAGAGGAGGTACAACC 500
Qy 792 -----IleGlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyr 807
Db 499 TACCAGACCTTAATTAAGCTGAAGGCTGGTGGTGAATAATGGAATCATACCAATATC 440
Qy 808 ArgLeuArgPheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeu 827
Db 439 TGGCTA -----GCTGCCAATGCGGCTC 419
Qy 828 SerProIleSerAspLysGlyTyr -----LeuThrGlyGlyGln 840
Db 418 TGTCATGCTAAATAATGATGGATATAATCTTCCTGGCATCACACATTTGACGCTGGCGGA 359
Qy 841 ValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAla 860
Db 358 -----AACAAACGCGACGAGGATCA 338
Qy 861 ValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGly 880
Db 337 CTGTATGCTGAATGGGGAACGTT -----GGAGCGTTTTCAGTAATTCGCAATTTACA 284
Qy 881 AlaGlyValGlyValArgTyrAlaSerProValGlyGlnValArgValAspValAlaThr 900
Db 283 CCG -----GGAGCTTACTGACAACT ----- 263
Qy 901 GlyValLysGluGluGlyAsnProIleLysLeuHisPhePheIle 915
Db 262 -----GAATCTGATGATTACAGTCGGCAGCTACTATGTG 230

RESULT 10

US-09-815-242-4760
; Sequence 4760, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4760
; LENGTH: 6228
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4760

Alignment Scores: 0.00432 Length: 6228
Pred. No.: 135.00 Matches: 154
Score:

Percent Similarity: 32.63% Conservative: 110
Best Local Similarity: 19.04% Mismatches: 301
Query Match: 2.86% Indels: 244
DB: 10 Gaps: 35
US-09-914-168-2 (1-919) x US-09-815-242-4760 (1-6228)
Qy 24 MetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValPro 43
Db 4144 GTTCTTCACAAGCATTAGCTAAAGAAAAAGAAAGACACTT ----- 4185
Qy 44 AlaHisaspThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrPro 63
Db 4186 -----GCAGCTATTGACCAAGCT ----- 4203
Qy 64 GluGlnIleGlnAlaArgLeuAsnAlaAlaGlyLeuAsn ----- 76
Db 4204 ---CAAACGAATAGTCAGGTGAATCAAGCAGCAACAATGTGTATCAGCGATTAAAAATT 4260
Qy 77 -----AlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspAspGln 92
Db 4261 ATTC AACCTGAAACAAAGTTAAACACGCTGCAGCTGCAAAAATCAAT --- 4308
Qy 93 SerProIleSerArgIleGlyGluGlnSerProProLeuGlyLeuAspMetSerValIle 112
Db 4309 -----CAAAAGCGCAATGAATACGTGCTAAGATTAAATCAAGTAAAGAACCAACA 4359
Qy 113 GluGluThrThrProLeuSerLeuGlu -----GluLeuPheAlaGlnGluSerThr 129
Db 4360 GCAGAAGAACAGTAGTACACTAGATAAAATCAATGAATTTGTAAATCAAGCCATGACA 4419
Qy 130 GluMetGlyIleAsn ---ProAsnAspTyrIleProGluTyrGlnGlyGluGlnProAsn 148
Db 4420 GATATTACGAATAATAGAACAAATCAACAAGTTGATGATACAAACAGTCAAGCGCTTGAT 4479
Qy 149 SerGluValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeu 168
Db 4480 AGCATTTGCTTTAGTG -----ACGCTGACCATATTGTTAGAGCA 4518
Qy 169 TyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyr 188
Db 4519 GCTGCTAGA -----GATGCGAGTTAAGCAA -----CAATATGAAGCTAAAGCGC 4563
Qy 189 GlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGluPro 208
Db 4564 GAAATTGAGCAAGCGGACATCGCACTCATGAAGAAAAACAAGTTGCTTTAAATCAATTA 4623
Qy 209 TyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAlaMetAspLeuAsn 228
Db 4624 GCGAATAATGAAAAACGTGCATTACAAAACATCGATCAAGCAATAGCG -----AAT 4674
Qy 229 GlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyr 248
Db 4675 AATGATGTGAACGCTGTTGAACAAATGCGATTGCTACACTAAAGGTGTA ----- 4725
Qy 249 AspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAsp 268
Db 4726 -----CAACCTCATATTGTAATTAAGCCT 4749
Qy 269 LeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGlyAlaAsp 288
Db 4750 -----GAAGCACAACAAGCAATAAAAGCAAGTGCAGAAAAATCAAGTGAATAAAAA 4803
Qy 289 AspLysAlaPheThrThrVal -----AlaAspGluValProLeuLeuIleGlyAspVal 306
Db 4804 GATACACCACATGCACACAGTTGATGAATAGTAGAAGCAATCAATTAATAGCCACACA 4863
Qy 307 PheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHis 326
Db 4864 CTCAAACAAGCGCAACAAGAA -----ATAGAAAAATCAAAATCAAGATGCT 4908
Qy 327 GlyTyrPheAspGlyArg ---TrpLeuAspArgSerValIleLeuProAsp --- 344
Db 344 -----ATAGAAAAATCAAAATCAAGATGCT 4908

```
Db 4909 GCTGTTACTGATGTTAGAAAATCAACAATCAAGCAATAGAGCAAAATAAACCTAAAGTA 4968
QY 345 -----AsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAsp 362
Db 4969 AGAGCTAAACGAGCTGGCTTGATAGATTGAAGAAATAATAAAT----- 5016
QY 363 GluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLys 382
Db 5017 -----CAACTCGATGCAATCCGAAATACGTTGGATACTACTCAAGATGAA 5061
QY 383 LeuProValLysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyr 402
Db 5062 AGAGATGTTGCTATTGATCTTTAATAAAT----- 5094
QY 403 AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMet 422
Db 5095 -----GTAATACAAATTAAATAATGACATTGCACAA-----AACAAA 5130
QY 423 ValAsnThrGluIleValPheProGluArgGluGlnIleGluAsnAspGlnVal----- 440
Db 5131 ACGAATGCAGAGTGCAGTCAACTGAGCTGATGGCAGCAGCAACATCAAGTGATTTTA 5190
QY 441 -----SerPheGluGlnSerSerSerSerArgThrGluPro--- 452
Db 5191 CCTAAGTTCAAGTTAAACGACGCGCTCAATCTGTTGGTGTAAAGCCGACGCTCAA 5250
QY 453 ---AlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAsp 471
Db 5251 AATGCACATCACTCAAGCAAGCTTTATCACTGAAGAAGAA----- 5292
QY 472 GlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLys 491
Db 5293 -----AGACTAGCTGCTAAACATTTAGTAGACAACAGCA 5325
QY 492 LeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeu 511
Db 5326 CTTATACAGGCTATTGATCAGATCAATCATCA-----GATAAGACTGCC 5370
QY 512 AlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaVal 531
Db 5371 CAAGTTAATCAAGATAGTATTAATGCTCAAAATATTTTCAAAATTAACCCAGCGACA 5430
QY 532 SerAlaValAlaArgAlaIle----- 538
Db 5431 ACAGTTAAGCAACAGCAGATCAACAATAATTCGCTACAAAATAAAATTAATTTA 5490
QY 539 -----LeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
Db 5491 ATTAAGCAAAATAACGAAGCGACAGATGAAGACAAATAATTCGAATAGCACAAGTTGAA 5550
QY 553 -----ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr 566
Db 5551 AAAGAGTTAATTAAGCTAAACACAAATTTGCTAGTGCAGTCACTAATGCAGATGTG--- 5607
QY 567 GlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGln 586
Db 5608 -----GCATATTTATTCGATGATGAGAAACAGAA----- 5637
QY 587 IleGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsn 606
Db 5638 -----ATTCGTGAATCGAA---CCT 5655
QY 607 LeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLys 626
Db 5656 GTTATTACAGAAGGGCTGCTCGAGAACAAATTGACAACATTTATTCACGATAAAAAA 5715
QY 627 GlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAla 646
Db 5716 -----CAAGCAATTGAAGCGAATATTCAAGCA 5742
QY 647 ThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThr 666
Db 5743 ACGGTAGAGAAAGAAATAGTATATTAGCACAGTTACAAATATTATGACACTGCTATT 5802
```

```
QY 667 ArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThr 686
Db 5803 GCACAAATTTGATCAAGATCGTAGC-----AATGCACAAATTTGATCAAAACA 5847
QY 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGln 706
Db 5848 GCATCATTTAAATCTA-----CAACAATACAT 5874
QY 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
Db 5875 GATTTAGATGTCATACCTCTATTAAA---AAGCCAGATGCTGAA----- 5913
QY 727 AlaValHisLysThrValAlaAlaAspAsnLeuValAsnPro-----MetArgGlyTyr 743
Db 5914 -----AAACGATTAATGATGATCTTGCACGCGCTACTGCTTTAGTSCAAAATAT 5964
QY 744 ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
Db 5965 CGAAAA-----GTAAGTAATCGTAATAGGCT 5991
QY 764 IleAlaArgAlaGlyIleSerGlyVal 772
Db 5992 GATGCATTTAAAGCTATAACTGCTTTA 6018

RESULT 11
US-09-815-242-8815
; Sequence 8815, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8815
; LENGTH: 6561
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(6561)
US-09-815-242-8815

Alignment Scores:
Pred. No.: 0.00467 Length: 6561
Score: 135.00 Matches: 154
Percent Similarity: 32.63% Conservative: 110
Best Local Similarity: 19.04% Mismatches: 301
```

```
Query Match: 2.86% Indels: 244
DB: 10 Gaps: 35
US-09-914-168-2 (1-919) x US-09-815-242-8815 (1-6561)

QY 24 MetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValPro 43
DB 4144 GTTGCTTCAACAGCATAGCTAGCTAAGAAAGAAAGAAAGCACTT----- 4185
QY 44 AlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyAsnProValLeuLeuThrPro 63
DB 4186 -----GCAGCTATTGACCAAGCT----- 4203
QY 64 GluGlnIleGlnAlaArgLeuAsnAlaAlaGlyLeuAsn----- 76
DB 4204 ---CAAAGCAATAGTCAGGTGAATCAAGCAGCAACAATGGTGTATCAGCGATTAAATTT 4260
QY 77 -----AlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspGln 92
DB 4261 ATTCACCTGAAACAAAAGTTAAACCCAGCTGCACGTGAATAAATCAAT----- 4308
QY 93 SerProIleSerArgIleGlyGlnGlnSerProProLeuGlyLeuAspMetSerValIle 112
DB 4309 -----CAAAACCGAATGANTACGTGTAGATTAATCAGGATAAAGCAACA 4359
QY 113 GluGluThrThrProLeuSerLeuGlu-----GluLeuPheAlaGlnGluSerThr 129
DB 4360 GCAGAAGAAAGACAAGTAGCAGCTAGATAAAATCAATGAATTTGTAAATCAAGCCATGACA 4419
QY 130 GluMetGlyIleAsn---ProAsnAspTyrIleProGluTyrGlnGlyGluGlnProAsn 148
DB 4420 GATATTACGAATAATAGAACAAATCAACAAAGTTGATGACAAACAGTCAACGCTTGAT 4479
QY 149 SerGluValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeu 168
DB 4480 ACCATTGCTTTAGTG-----ACCGCTGACCATATTGTAGAGCA 4518
QY 169 TyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyr 188
DB 4519 GCTGCTAGA-----GATCAGTTAAGCAA-----CAATATGAAGCTAAAGCGC 4563
QY 189 GlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGluPro 208
DB 4564 GAAATTGACCAAGCGGAACATCGGACTGATGAAGAAACAAAGTTCGTTTAATCAATTA 4623
QY 209 TyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGlnSerAlaMetAspLeuAsn 228
DB 4624 GCGAATTAATGAAAACGTCATTACAAAACATCGATCAAGCAATAGCG-----AAT 4674
QY 229 GlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyr 248
DB 4675 AATGATGTGAACGCTGTTGAAAACAATGTCATTGTCTACACTAAAGGTGTA----- 4725
QY 249 AspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAsp 268
DB 4726 -----CAACCTCATATTGTAATTAAGCCT 4749
QY 269 LeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGlyAlaAsp 288
DB 4750 -----GAAGCACACAACCAATAAAGCAAGTCGAGAAAATCAAGTCAATCAATAAA 4803
QY 289 AspLysAlaPheThrThrVal-----AlaAspGluValProLeuLeuIleGlyAspVal 306
DB 4804 GATACACCACATGCAACAGCTTGATGAATTAGATGAAGCGCAATCAATTAATTAGCGCACAC 4863
QY 307 PheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHis 326
DB 4864 CTCAAACAAGCCACAACAG-----ATAGAAAATCAAAATCAAGATGCT 4908
QY 327 GlyTyrPheAspGlyArg---TrpLeuAspArgSerValAspValIleLeuProAsp--- 344
DB 4909 GCTGTTACTGATGTTAGAAATCAACAAATCAAGGCAATAGACCAATAAAACCTTAAGTA 4968
```

```
QY 345 -----AsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAsp 362
DB 4969 AGACGTAAACGAGCTGCGCTTGCATGATGAAAGAAATAATAAAAT----- 5016
QY 363 GluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLys 382
DB 5017 -----CAACTCGATCAATCCGAAATACGCTTGATCTACTACTCAAGATGAA 5061
QY 383 LeuProValLysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGlyAlaTyr 402
DB 5062 AGAGATGTTGCTATTGATACATTTAAATAAAT----- 5094
QY 403 AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMet 422
DB 5095 -----GTAAATACAATTAATAATGACATTGCACAA-----AACAAA 5130
QY 423 ValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnVal----- 440
DB 5131 ACGAATGCAGAAAGTGATGAACTGAGACTGATGCAACGACACATCAAGATGATTTTA 5190
QY 441 -----SerPheGluGlnSerSerSerSerSerArgThrGluPro--- 452
DB 5191 CCTAAAGTTCAAGTTAAACCCAGCAGCGCTCAATCTGTTGGTGAAGCCGCAAGCTCAA 5250
QY 453 ---AlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAsp 471
DB 5251 AATGCACATAATCGATCAAGCGATTATCACTCAAGAAAGAA----- 5292
QY 472 GlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLys 491
DB 5293 -----AGACTAGCTGTAAACATTTAGTAAACAAGCA 5325
QY 492 LeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspArgValLeu 511
DB 5326 CTTAATCAGCGCTATTGATCAGATCAATCATGCA-----CATAGACTGCC 5370
QY 512 AlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaVal 531
DB 5371 CAAGTTAATCAAGATAGTATAAATGCTCAAAATATATTATTCAAAAATTAACACGCGACA 5430
QY 532 SerAlaValAlaArgAlaIle----- 538
DB 5431 ACAGTTAAAGCAACAGCATTTACAACAAATTCACAAATATCGCTACAAATATAATTTA 5490
QY 539 -----LeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
DB 5491 ATTAAGCAAAATAACGACGACAGATGAAGAACAAAATATTGCAATAGCACAAAGTTGAA 5550
QY 553 -----ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr 566
DB 5551 AAAGAGTTAATTAAGCTAAACCAACAAATTTGCTAGTGCAGTACTAATCGAGATGTG--- 5607
QY 567 GlnSerLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGln 586
DB 5608 -----CCATATTATTGTCATGTCAGAAAACGAA----- 5637
QY 587 IleGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsn 606
DB 5638 -----ATTCGTCAAAATCGAA---CCT 5655
QY 607 LeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLys 626
DB 5656 GTTATTAAACAGAAAGCGCTCTGCTCGAGAACAAATTTGACAACATATTATTCAACGATAAANA 5715
QY 627 GlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAla 646
DB 5716 -----CAAGCAATTTGAACGCGAATATTTCAGCA 5742
QY 647 ThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThr 666
DB 5743 ACGGTAGAGAAGAAATAGTATTATAGCACAGTTACAAAATATTATGACACTGCTATT 5802
QY 667 ArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThr 686
```

Db 5803 GGACAAATTCATCAAGATCGTAGC-----AATGCACAACTTCATAAAACA 5847
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGln 706
Db 5848 GCATCATTAATCTA-----CAAACAATACAT 5874
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
Db 5875 GATTTAGATGATACCTCTATTAA-----AAGCCAGATGCTGAA----- 5913
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnPro-----MetArgGlyTyr 743
Db 5914 -----AAACGATTAATGATCTGCGCGCTGCTGCTTTAGTGCAAAATAT 5964
Qy 744 ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
Db 5965 CCAAAA-----GTAACTATCGTAAAGGCT 5991
Qy 764 IleAlaArgAlaGlyIleSerGlyVal 772
Db 5992 GATGCATTAAGCTATAACTGCTTTA 6018

RESULT 12

US-09-765-272-117
; Sequence 117, Application US/09765272
; Patent No. US20020061545A1
; GENERAL INFORMATION:
; APPLICANT: Choi et. al.
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
; NUMBER OF SEQUENCES: 452
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09765,272
; FILING DATE: 22-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/961,083
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 117:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3121 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 117:

US-09-765-272-117

Alignment Scores:
Pred. No.: 0.00198 Length: 3121
Score: 134.00 Matches: 190
Percent Similarity: 32.04% Conservative: 124
Best Local Similarity: 19.39% Mismatches: 318
Query Match: 2.83% Indels: 348
DB: 10 Gaps: 53

US-09-914-168-2 (1-919) x US-09-765-272-117 (1-3121)
Qy 24 MetThrSerGlnAlaLeuAlaGlnGlnAsnProAlaAsnIle-----IleAsnHis 41
Db 155 TTGACCAAGCCGATTTTATCTGCTATAAATAGTCAGCTTTCTATCGGAGTCGGGCAACAT 214
Qy 42 ValPro-----AlaHisAspThrAlaIleAsnGlnAlaLys 53
Db 215 TTACCAAGAGCCCTTGAAGGTTATCAATATATATGTTATATCAAACTAACAA 274
Qy 54 AlaGlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaArgLeuAsnAlaAla 73
Db 275 CAGGATATACAGAGCTTTCAAGGACA-----GTTGATGGGAAATCTCTGCTCAA 325
Qy 74 GlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspAsp----- 91
Db 326 AGAGATAGTCAACCAAACTCTACAAAAAATCAGATCTAGTTTCATTTCAGCTGATTTAGAA 385
Qy 92 -----GlnSerProIleSerArgIleGlyGluGlnSerProProLeuGlyLeu 107
Db 386 TGGAAACCAAGCAGAGGGAAGGTTAGTTTACAAGGTGAAGCATCAGGGGATGATGCACTT 445
Qy 108 AspMetSerValIleGluGluThrThrProLeuSerLeuGluLeuPheAlaGlnGlu 127
Db 446 -----TCAGAAAAATCTTCTATAGCAGCAGACAATCTATCTCTTAATGAT 490
Qy 128 SerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnPro 147
Db 491 TCATTCCCAAGTCAAGTTGAGCAGAAT-----CCGGATCACAAGAGGAATCT--- 538
Qy 148 AsnSerGluValValProThrLeuGluProGluLysProGlyLeuIleLysArg 167
Db 539 -----GTAGTTCGACCAACAGTG---CCAGACAAGAAATCTCTGTCTGCT 583
Qy 168 LeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPhe 187
Db 583 ----- 583
Qy 188 TyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGlu 207
Db 584 ---ACAACGGTCGAGAGTCGGGAAGAGGATTTGCGCAGCACAATCATCGACCCAG 640
Qy 208 ProTyrAlaAsnIleLysAlaLeuGluAspIle---ThrGlnGluSerAlaMetAsp 226
Db 641 -----TATAAACTTCCATTGGAAACCAAGCAGCAGCAAGCCGCTCATGAG 688
Qy 227 LeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGly 246
Db 688 ----- 688
Qy 247 TyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIle 266
Db 689 -----GGTGAAGCC---CGAGTCCGT 706
Qy 267 HisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGly---Glu 285
Db 707 GAAGACTTA-----CCAGTCTACACT-----AAGCCACTAGAAACCAAGCTACACAA 754
Qy 286 GlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAsp 305
Db 755 GGACCCGGACATGAAGGTGAAGCTGCAGTTCCGGAAGAAACACAGCTTACACAAACCG 814
Qy 306 ValPhe-----HisHisGlyLysTyrGluThrLysLysAsn 317
Db 815 TTAGCAACGAAAGCAGCAGCAGCAGCTCATGAGGCAAGACTACAGTCCGCCAAGAG 874
Qy 318 LeuIleGlu-----AsnAlaSerAlaGluHisGlyTyr 328
Db 875 ACTCTAGATACACCGAACCGGTACCGACAAAGGCACACAAAGACCCGAAACAT----- 928
Qy 329 PheAspGlyArgTrpLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAsp 348
Db ----- 348

Db 929 ---GAGGC-----GAACGSCAGTAGAAGAAGAACTTCG----- 961
Qy 349 ValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThr 368
Db 962 ---GCTTTAGGTCACACGAAATAGAACGGAATCCAGAAATAT----- 1006
Qy 369 IleAspProLysThrAsnGlnLeuThrAspProAspLysLeuProValLysArgGlu 388
Db 1007 ---CCTTATACAACAGAAGAAATTCAGGATCCACACTTCTGAAAAATCGTCGTAAG 1060
Qy 389 LeuLeuGluGlnLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArg 408
Db 1061 ATTGAACGACAA-----GGCAACGA----- 1081
Qy 409 AlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsn-----MetValAsn 424
Db 1082 -----GGGACACGTCACAAATTCAAATATGAGACTACATCGTAAAT 1120
Qy 425 ThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGln 444
Db 1121 GGTAAATGTCGTA-----GAAACT 1138
Qy 445 SerSerSerArgThrGluProAlaGlnValAspGlu-----SerThrLeu 460
Db 1139 AAGAAGTGTACCAAGTCAAGTAGCTCCGGTCAACGAAGCTGTTAAAGTAGGAACACTT 1198
Qy 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIle 480
Db 1199 GTGAAAGTTAAACCTACAGTAGAATAATACA-----AACTTAACAAGATT 1243
Qy 481 GluPheSerAlaSer-----AsnLeuIleGlnAspLysLeuAsnLeuVal 495
Db 1244 GAGAACAAAAATCTATAACTGTAAAGTTAATACTTAATAGACACTACCTCAGCATATGTT 1303
Qy 496 AlaAlaLysAlaArgHisLeuTyrAspMetProAspArgValLeuAlaIleAsnHis 515
Db 1304 TCTCAAAAACGCAA-----GTTTTCCTAT 1327
Qy 516 AspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAla 535
Db 1328 GGAGAC-----AAGCTAGTTAAA 1345
Qy 536 ArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAla 555
Db 1346 GAGGTGGATATAGAAAATCTGCGCAAGAGCAAGTAATA----- 1384
Qy 556 LeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeuTyrVal 575
Db 1385 -----TCAGGTTTAGATTACTACACACCGTATACAGTTAAACACACCTA 1429
Qy 576 PheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThr 595
Db 1430 -----ACTTATAATTTGGTGAAATAATGAGGAA 1459
Qy 596 GlyThrArgLeuValThr-----LysPheGluHisAsnLeuIleAsnArgAspGly 612
Db 1460 AATACTGAACATCAACTCAAGATTTCCAAATTAGATATAGAAAATAGAG----- 1510
Qy 613 TyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeuTyrAla 632
Db 1511 -----ATTAAGATATTGATTCAGTAGAATATACCGT 1543
Qy 633 ThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGln 652
Db 1544 AAGAAA-----AATGATCGTTATCGTAGATATTAAGTCTAAGT--- 1582
Qy 653 GluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGlu 672
Db 1583 -----GAAGCGCGGACTGATCGGCTAAA----- 1606
Qy 673 IleSerArgSerIleIleGlnAsnGlyGlyTyrPheAsnArgThrTyrSerLeuArgTyrArg 692
Db 1607 -----TACTTTGTAAGAGTGAAG 1624

Qy 693 LeuAspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPhe 712
Db 1625 TCAGATCGCTTCAAGAAATGTAC-----CTACCTGTAAATCT 1663
Qy 713 ValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrVal 732
Db 1664 ATTACA---GAAAATACGGATGGAAACGTATAAAGTAGCGGTAGCGTTGATCAACTGTCT 1720
Qy 733 AlaAspAsnLeuValAsnProMetArgLysTyrArgGlnArgTyrSerLeuGluValGly 752
Db 1721 GAAGAAGGT-----ACAGACGGTTACAAAGATGATTACACATTTACTGTAGCT 1768
Qy 753 SerSer-----GlyLeuValSerAspAlaAsnMetAlaIleAla----- 765
Db 1769 AATCTAAACGAGACCAACAGAGATTACACATCCTTTAAACAGCTGGTAAACGCCATG 1828
Qy 766 ArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsn-----Ala 779
Db 1829 CAAAGCAATCTGTCTGTCTATACATTTGGCTTCAGATATGACCGCAGATGAGTGAGC 1888
Qy 780 TyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSer 799
Db 1889 TTAGCGCATAGACAGCAAGTTATCTCACAGGTGCATTTACAGGAGCTTGTGCGTTCT 1948
Qy 800 AspAsnPhe---AsnHisValProTyrArgLeuArg-----PhePheAla 813
Db 1949 GATGAACAAAATCGTATGCCATTTATGATTTGAAGAAACCATTTATTGATACATTAAT 2008
Qy 814 GlyGly-----AspGlnSerIleArgGlyTyrAlaHisAspSer----- 826
Db 2009 GGTCTACAGTTAGAGATTTGGATATTAATAACTGTTCTGCTGATAGTAAAGAAATGTC 2068
Qy 827 -----LeuSerProIleSerAspLysGlyTyr 835
Db 2069 GCACCGTGGCGAAGGACGCAATAGCCGGAATATTATATGTTGCGTAGAAGGAAAA 2128
Qy 836 LeuThrGlyGlyGlnValLeu-----AlaValGlyThrAlaGluTyr 849
Db 2129 ATCTCAGGTGCGAAATCTGTTGCGGATTAGTAGCGAGCGCAACAATACAGTGATGAA 2188
Qy 850 AsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr 869
Db 2189 AACAGCTCGTTTACAGGGAACCTT-----ATCGCAAAATCACCCAG 2227
Qy 870 AspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGly-----ValArgTrp 887
Db 2228 GACAGTAAAT---AAAAATGATAC---GGAGGAATAGTAGGTAAATATACAGGAAAT 2278
Qy 888 AlaSerProValGlyGlnValArgValAspValAlaThrGlyValLysGluGlyAsn 907
Db 2279 AGTTGAGAGTTAATAAGTTAGGTAGATGCCCTTAATCTCTACTAATGACGCAATAAT 2338

RESULT 13

US-10-160-758-6
; Sequence 6, Application US/10160758
; Publication No. US20030036076A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: CADS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-089C
; CURRENT APPLICATION NUMBER: US/10/160,758
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 14536

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-160-758-6

Alignment Scores:
Pred. No.: 0.0211 Length: 14536
Score: 133.50 Matches: 199
Percent Similarity: 31.69% Conservative: 162
Best Local Similarity: 17.47% Mismatches: 393
Query Match: 2.82% Indels: 385
DB: 9 Gaps: 50

US-09-914-168-2 (1-919) x US-10-160-758-6 (1-14536)

QY 52 AlAlaAlaGlyAsnProProValLeuLeuThrProGluGln-----lLeGlnAlaArg 69
   |||:|:| ||| ||| ||| |||:|:| |||:|:| |||:|:| |||:|:| |||:|:|
Db 6446 GCTCGGATGGAGAAAGCCATCCCTCCAGAGTGAGAGAGGTACTTGTCACGTGTGAGA 6505

QY 70 LeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPhe 89
   |||:|:| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6506 AATAAATCCAAACCCACTGTTTCAGAGTCCCTATTACAAAGTCAGAGTACCTGAAATATC 6565

QY 90 AspAspGlnSerProIleSerArgIleGlyGluGlnSerProPro---LeuGlyLeuAsp 108
   |||:|:| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6566 ACCCTCTATACCCCAATTCTCCACACCCAGGCCCGGAGTCCAGAGGACTCCGGCTCATC 6625

QY 109 MetSerValIleGluGlu-----|-----|-----|-----|----- 114
   |||:|:| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6626 TACAACATTTGGAGGAAGAACCTTGATGCTTTCACCACTGACTTCAAGACTGGTGTC 6685

QY 115 -----ThrThrProLeuSerLeuGlu-----|-----|-----|-----|----- 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6686 CTAACAGTAAACGGGCTTTGGACTATGATGCCAAGACCAACATGTGTTCAACAGTCAGA 6745

QY 128 SerThrGluMetGlyIleAsnPro---AsnAspTyrIleProGluTyrGlnGlyGluGln 146
   |||:|:| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6746 GCCACGGATACAGCTCTGGGGTCATTTCTGAAGCCACAGTGGAGTCTAGTGGAGGAT 6805

QY 147 ProAsnSerGluValValProProThrLeuGlu-----|-----|-----|----- 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6806 GTCATATGATAAC-----|-----|-----|-----|-----|----- 164

QY 159 -----ProGluLysProGlyLeuIleLysArgLeuTyrAla-----|----- 170
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6857 TCAGAAGGCTTGCTCTCAGACCCCTGTGATCCCACTGTGCTCTCAGCAGGACTCA 6916

QY 171 -----|-----|-----|-----|-----|-----|-----|----- 6964
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6917 GGCGGAACCGTGACGCTCTTATCATGATTTGGAGATTTGGAGATGGCTCATGAT 6964

QY 183 LeuLysAlaLysPheTyrGln---SerSerGlnSerGlyGluThrSerAlaIleGlySer 201
   |||:|:| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6965 ---GTTTCCAAAGTCTTCAGATCAATGGGACGACAGGGGAGATGTCACAGTTCGAAGAA 7021

QY 202 SerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp----- 218
   |||:|:| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7022 CTGGATTATGAAGCCCAACACACATTTTCATGTGAAGTCAAGGATGATGAAGAGAGAT 7081

QY 219 -----|-----|-----|-----|-----|-----|-----|----- 231
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7082 CCCCCACTCACTGGTGAACCCCTTGTTGGTTGTCAATGTGCTGATATCAATGACAACCC 7141

QY 232 ProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrAspIleAsp 251
   |||:|:| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7142 CCAGATTCCAGACACCTCAATATGAGCCCAATGTCAGT-----|-----|-----|----- 7183

QY 252 LeuSerIleAArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGlu 271
   |||:|:| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7184 CTGGCAACCTGTGGACACTTGTTTAAAGTCCAGGCTATTGACCCCTGACAGCAGAGAC 7243

QY 272 ProValTyrIleAspTyrArgAlaValGluValArgGlyGlu-----|----- 285
   |||:|:| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7244 ACCTCCCGCTGGAGTACCTGATTTCTTTCGCAATCAGCAGGACGACTTCTTTCATTAAAC 7303
```

```
Qy 563 AlaAsp-----ValTyrGln----- 567
    |||
Db 8276 ATGACCACGAATCCACCAAAATTGTACCAGATTGATGTGGCAGCATTTGCCTTCAGAAC 8335
Qy 567 ----- 567
Db 8336 ACTGATGTGGTGCTCTGGTCTCTGTCAACATCCAAATGGGAGACGTCATATGACAATAGG 8395
Qy 568 -----SerLysLysValProLeu----- 573
    |||
Db 8396 CCTGTATTGAGCGTGATCCATATAAGCGTCTCTCACTGAGATAATGCCAGTGGGGACC 8455
Qy 574 -----TyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeu 589
    |||
Db 8456 TCAGTCATTCAAAGTGACTGGCATTTGACAAGGACACTGGGAGAGATGGCCAGGTGAGCTAC 8515
Qy 590 GlyTrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsn 609
    |||
Db 8516 AGGCTGTCTGCAGACCCTGGTAGCAATGTC-----CATGAGCTCTTTGCC 8560
Qy 610 ArgAsp----- 611
    |||
Db 8561 ATTGACAGTGAGAGTGGTGGATCACCACACTCCAGGAACCTTGACTGTGAGACCTGCCAG 8620
Qy 612 -----GlyTyrGlnAlaGlyAlaGluLeuArgLeuSer----- 622
    |||
Db 8621 ACTTATCATTTTCATGTGTGGTCTATGACCACGACAGACCATCCAGCTATCCTCTCAG 8680
Qy 622 ----- 622
Db 8681 GCCCTGTTTCAGGTCTCCATTACAGATGAGAAATGCTCCCGATTGCTTCTGAA 8740
Qy 623 -----GluAspLysLysGlyValLysLeuTyrAlaThr---Lys 634
    |||
Db 8741 GAGTACAGAGGATCTGTGGTTGAGAACAGTCAGCGCTGGCAACTGGTGGCACTCTAAG 8800
Qy 635 ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGlu--- 653
    |||
Db 8801 ACCCTGGATGCTGACATTTCTTGACAGAACAGGCGAGGTCACCTGCTACATCAGAGAGGA 8860
Qy 654 -----ValPheGlyHisSerThrAsnGlyPheAsp-----LeuSerThr 666
    |||
Db 8861 GACCCCTGGCCAGTTTGGCATCAGCCAGTTGGAGATGAGTGAGGATTTCTCTCAAG 8920
Qy 667 ArgThrLeuGluHisGlyIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThr 686
    |||
Db 8921 AAGACCTGGACCGCAGCATACAGCAAG----- 8950
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGln 706
    |||
Db 8951 TACTTGTCTCAGAGTCACAGCATCTGTATGGCAAGTTCAGGCT---TCGGTCACCTGTGGAG 9007
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal 726
    |||
Db 9008 ATCTTTGCTGGACGTCATATGATAACAGCCACAGTGTTCACAGCTTCTCTATACTGGC 9067
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
    |||
Db 9068 AAGTTTCATGAAGATGTA-----TTTCCAGGACAC 9097
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArg 766
    |||
Db 9098 TTCATTTTGAAGGTTTCTGCCACAGACTTGGACACTGATACCAATGTCTCAGATC----- 9151
Qy 767 AlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHis 786
    |||
Db 9152 -----ACATATTTCTCTG-----CATGCCCTTGGG---GCGCAT 9181
Qy 787 GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSer----- 799
    |||
Db 9182 GAATTCAAGCTGGATCTCATACAGGGGAGCTGACACACTCAGCTGCCCTAGACCGAGAA 9241
Qy 800 -----AspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGln 817
```

```
Db 9242 AGGAAGGATGTGTTCAACCTTGTT-----GCCAAGGCGAGGATGGAGGTGGCGGA 9292
Qy 818 SerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThr 837
    |||
Db 9293 TCGTGGCAGGAGACATCACCTCCATGTGGAGGATGTGAATGACAATGCC----- 9343
Qy 838 GlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeu 857
    |||
Db 9344 -----CCGCGGTTCTTCCCCAGGCACCTGT 9367
Qy 858 ArgLeuAlaValPheGlyAsp-----IleGlyAsnAlaTyrAspLysGlyPheThrAsn 875
    |||
Db 9368 GCTGTGCTGTCTCGACACACACACAGTGAAGACCCCTGTGTGGCTGTAGTATTTGCCCGG 9427
Qy 876 AspThrLysIleGlyAlaGlyValGlyValArgTrpAlaSerPro-----ValGly 892
    |||
Db 9428 GATCCCCAGCAAGCGGCAATGCCAGGTGTTTACTCTCTGCGGATTCACCCGAAGGC 9487
Qy 893 GlnValArgValAspValAlaThrGlyValLysGluGluGlyAsnProIleLysLeu 911
    |||
Db 9488 CACTTTTCATCGACGCCACCCAGCGGGTGATCCGCTGGAAGAACCCGCTGCAGGTC 9544
```

RESULT 14

US-10-160-758-7

; Sequence 7, Application US/10160758

; Publication No. US20030036076A1

; GENERAL INFORMATION:

; APPLICANT: EXELIXIS, INC.

; TITLE OF INVENTION: CADS AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE

; FILE REFERENCE: EX02-089C

; CURRENT APPLICATION NUMBER: US/10/160,758

; PRIOR FILING DATE: 2002-06-03

; PRIOR APPLICATION NUMBER: US 60/296,076

; PRIOR FILING DATE: 2001-06-05

; PRIOR APPLICATION NUMBER: US 60/328,605

; PRIOR FILING DATE: 2001-10-10

; PRIOR APPLICATION NUMBER: US 60/357,253

; PRIOR FILING DATE: 2002-02-15

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 7

; LENGTH: 14536

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-10-160-758-7

Alignment Scores:

Pred. No.: 0.0211 Length: 14536

Score: 133.50 Matches: 199

Percent Similarity: 31.69% Conservative: 162

Best Local Similarity: 17.47% Mismatches: 393

Query Match: 2.82% Indels: 385

DB: 9 Gaps: 50

US-09-914-168-2 (1-919) x US-10-160-758-7 (1-14536)

Qy 52 AlaLysAlaGlyAsnProProValLeuLeuThrProGluGln-----IleGlnAlaArg 69

Db 6446 GCTCGGATGGAGAACCCCATCTCCAGAGTGAAGAGGACTTGTCTACTGTGAGA 6505

Qy 70 LeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPhe 89

Db 6506 AATAATCCAAACCCACTGTTTTCAGAGTCCTTATTACAAAGTCAGAGTACCTGAAATATC 6565

Qy 90 AspAspGlnSerProIleSerArgIleGlyGlnSerProPro---LeuGlyLeuAsp 108

Db 6566 ACCCTCTATACCCCAATTCACACACCCAGCCCGGAGTCCAGAGGACTCCGGCTCATC 6625

Qy 109 MetSerValIleGluGlu----- 114

Db 6626 TACAACATGTGGAGGAAGAACCCTTGATGCTGTTCACCACTGACTTCAAGACTGGTCTC 6685

QY 115 -----ThrThrProLeuSerLeuGlu-----GluLeuPheAlaGlnGlu 127
Db 6686 CTAACAGTAACAGGGCTTTGGACTATGAGTCCAAAGACCAACAAATGTTTCACAGTCAGA 6745
QY 128 SerThrGluMetGlyIleAsnPro---AsnAspTyrIleProGluTyrGlnGlyGluGln 146
Db 6746 GCCACGATACAGCTCTGGGGTCATTTCTGAAGCCACAGTGGAAGTCCCTAGTGGAGGAT 6805
QY 147 ProAsnSerGluValValProProThrLeuGlu----- 158
Db 6806 GTCAATGATAAC-----CCTCCCACTTTTCCCAATTTGGTCTATACCACTTCATC 6856
QY 159 -----ProGluLysProGlyLeuIleLysArgLeuTyrAla----- 170
Db 6857 TCAGAGGCTTGCTGCTCAGACCCCTGTGATCCCACTGTGTGCTCTGTGACCAGGACTCA 6916
QY 171 -----ArgLeuPheAsnAspGlyValAlaAsnLysValProArg 182
Db 6917 GGGCGGAACCGTCAGCTCTCTTATCAGATTGTGGAGGATGGCTCAGAT----- 6964
QY 183 LeuLysAlaLysPheTyrGln---SerSerGlnSerGlyGluThrSerAlaIleGlySer 201
Db 6965 ---GTTTCCAAGTCTTCCAGATCAATGGGAGCAGACAGGGGAGATGTCCACAGTTCAAGAA 7021
QY 202 SerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp----- 218
Db 7022 CTGGATTAGAACCCCAACACACTTTCATGTGAAAAGTCAGGCCCATGATAAAGGAGAT 7081
QY 219 -----IleThrGlnGluSerAlaMet-----AspLeuAsnGlySerIle 231
Db 7082 CCCCCACTCACTGGTGAACCCCTGGTGTCAATGTGCTGATATCAATGACACACCCC 7141
QY 232 ProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAsp 251
Db 7142 CCAGAGTTCAAGACACCTCAATATGAAGCCATGTCACT-----GAA 7183
QY 252 LeuSerIleIleArgAsnSerIleGlyValValAspValIleIleHisAspLeuGlyGlu 271
Db 7184 CTGGCAACCTGTGGACACCTGGTCTTAAAGTCCAGGCTATTGACCTCAGCAGCAGAC 7243
QY 272 ProValTyrIleAspTyrArgAlaValGluValArgGlyGlu----- 285
Db 7244 ACCTCCCGCTGGAGTACCTGATCTTCTTGCAATCAGGAGGCACTTCTTCATTAAC 7303
QY 285 ----- 285
Db 7304 AGCTCATCGGGAATAATTTCTATGTTCAACCTTTGCAAAAAGCACCTGGACTCTCTTAC 7363
QY 286 -----GlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeu 301
Db 7364 AATTTGAGGATAGGTGCTTCTGATGGAGTCTCCGAGCAACT-----GTGCGCTGG 7414
QY 302 LeuIleGlyAspVal-----PheHisHisGlyLysTyrGlu 313
Db 7415 TACATCAACACTACAAATCCCAACAGTACAGCCAGAGTTCACGAGCACCCTTTATGAG 7474
QY 314 ThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgIrp 333
Db 7475 GCAGAA-----TTAGCAGAGATGCAATGGTGGAAACCAAGGTGATTGTGTAGCC 7528
QY 334 LeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyr 353
Db 7529 ATAGACAAAGATAGTGGT-----CCCTATGGCACTATAGATATATCATCATCAAT 7579
QY 354 AspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIleAspProLysThr 373
Db 7580 AAACGTAGCAAGTGAAGA-----TTCTCCATAAAGCCC---AAT 7615
QY 374 AsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGluGlnLeu 393
Db 7616 GCCAGATTGCCACT-----CTCAGAAACTG 7642
QY 394 LeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAsp 413

Db 7643 GATCGGGAATAATTCAACAGAG-----AGAGTCATTGCTATTAAAGTCATGGCTCGGGAT 7696
QY 413 ----- 413
Db 7697 GGAGGAGGAAGAGTAGCTTCTGCACGGTGAAGATCATCTCACAGATGAAATATGACAAC 7756
QY 414 -----LeuIleAlaThrArgTyrPheAsnMetValAsnThr----- 425
Db 7757 CCCCCACAGTTCAACAGCATCTGAGTACACAGTATCCATTCAATCCCAATGTCAGTAAAGAC 7816
QY 426 -----GluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSer 441
Db 7817 TCTCCGGTTATCCAGGTGTGGCTATGATGATGATGAAGGTGAGAGTGCAGACCCAGATGTCCACC 7876
QY 442 PheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGlu 461
Db 7877 TAC-----TCAGTGAACCCAGAGGACCTAGTTAAAGATGTCTATTCGAATTAAC 7924
QY 462 ProValIleGluThrValGluLeuThrAspGlyIleLeu---MetAspIleSerProIle 480
Db 7925 CCAGTCACCTGGTGTGTCGAAGGTGAAGACACGCTGGTGGGATTGGAATATCAGACCCTT 7984
QY 481 GluPheSerAla-----SerAsnLeuIleGlnAsp 490
Db 7985 GACTTCTTCATCAAGCCCAAGATGGAGGCCCTCCTCACTGGAACCTCTCTGTCGCCAGTA 8044
QY 491 LysLeuAsnLeuValAlaAlaLysAlaArg-----HisLeu 502
Db 8045 CGACTTCAGTGGTCTCTTAAAGATATCCTTACCGAAATTTTCTGCAACCTTTGTACTACT 8104
QY 503 TyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSer 522
Db 8105 TTCTCTGCACCTGAAGAC-----CTTCCAGAGGGGTCTGAAATTTGGGATTCTTAAAGCA 8158
QY 523 IleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGlu 542
Db 8159 GTGGCAGCTCAA---GATCCAGTCACTACAGTCTAGTGGGGGCACTACACCTCAGAGC 8215
QY 543 SerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrPro 562
Db 8216 AACAAGGATGGTGTCTCTCCCTAGACCCACAGACACAGGGGTCTAAAGGTGAGGAAGCCC 8275
QY 563 AlaAsp-----ValTyrGln----- 567
Db 8276 ATGGACCACGAATCCACAAATTTGACCAGATTGATGTGATGGCACATTTGCCCTTCAGAAC 8335
QY 567 ----- 567
Db 8336 ACTGATGTGCTCTCTCTGCTCTCAACATCCAAAGTGGGAGGAGTCAATGACAAATAGG 8395
QY 568 -----SerLysLysValProLeu----- 573
Db 8396 CCTGTATTGAGGCTGATCCATATAAGGCTGTCTCTCACTGAGATATGCCAGTGGGACC 8455
QY 574 -----TyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeu 589
Db 8456 TCAGTCATTCAAGTGACTGCTCATTTGACAAAGGACACTGGGAGAGATGGCCAGGTACAGTAC 8515
QY 590 GlyTrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsn 609
Db 8516 AGGTGTCTGCAGACCCCTGGTAGCAATGTC-----CATGAGCTCTTTGCC 8560
QY 610 ArgAsp----- 611
Db 8561 ATTGACAGTGAGAGTGGTTGGATCACCACACTCCAGGAACCTTGACTGTGAGACCTGCCAG 8620
QY 612 -----GlyTyrGlnAlaGlyAlaGluLeuArgLeuSer----- 622
Db 8621 ACTTATCATTTTTCATGTGTGGCTTATGACCACGAGACCACTCCAGTATCTCTCTCAG 8680
QY 622 ----- 622

```
Db 8681 GCCCTGGTTCCAGTCTCCATTACAGATGAGATGACAATGCTCCCGATTGCTTCTTGAA 8740
Qy 623 -----GluAspLysLysGlyValLysLeuTyrAlaThr---Lys 634
Db 8741 GAGTACAGAGGATCTGGTTGAGAACAGTGGCCGAACTGGTGGGACTCTAAG 8800
Qy 635 ProLeuSerHisProLeuAsnAspGlnLeuAlaThrLeuGlyTyrGlnGlnGlu--- 653
Db 8801 ACCCTGGATGCTGACATTCTGAGCAGAACAGGAGGTCACCTGTACATCACAGAGGA 8860
Qy 654 -----ValPheGlyHisSerThrAsnGlyPheAsp-----LeuSerThr 666
Db 8861 GACCCCTGGCCAGCTTGGCATGACCAAGTTGGAGATGGAGGAGATTTCCTCAGG 8920
Qy 667 ArgThrLeuGluHisGluLeuSerArgSerIleIleGlnAsnGlyTyrAsnArgThr 686
Db 8921 AAGACCTGGCCGCGAGCATACGCCAAG----- 8950
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTrpGln 706
Db 8951 TACTTGTCTGGAGCTCAATGATACAGCCACAGTGTTCACAGCTCTCTATCTACTGC 9007
Qy 707 AspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGlyVal 726
Db 9008 ATCTTGTCTGGAGCTCAATGATACAGCCACAGTGTTCACAGCTCTCTATCTACTGC 9067
Qy 727 AlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArg 746
Db 9068 AAGTTTCAAGATGTA-----TTTCCAGGACAC 9097
Qy 747 TyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArg 766
Db 9098 TTCATTTTGAAGTTTCTGCCACAGACTTGGACACTGATACCAATGCTCAGATC----- 9151
Qy 767 AlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHis 786
Db 9152 -----ACATATCTCTG-----CATGGCCCTGGG---CGCAT 9181
Qy 787 GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrpSer----- 799
Db 9182 GAATTCAGAGCTGGATCTCATACAGGGAGCTGACACACTCAGTCCCTAGACCGAGAA 9241
Qy 800 -----AspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGln 817
Db 9242 AGGAAGATGTGTTCAACTTGT-----GCCAAGGGCAGCGATGGAGGTGGCCGA 9292
Qy 818 SerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThr 837
Db 9293 TCGTCCAGGCGAGCATCACCTCCATGTGGAGGATGGAATGACAATGCC----- 9343
Qy 838 GlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMetLysAspLeu 857
Db 9344 -----CCGCGGTCTCTCCCGAGCCACTGT 9367
Qy 858 ArgLeuAlaValPheGlyAsp-----IleGlyAsnAlaTyrAspLysPheThrAsn 875
Db 9368 GCTGTGGCTGTCTTCGAACAACACAGAGTGAAGACCCCTGTGGCTGTAGTATTGCCCGG 9427
Qy 876 AspThrLysIleGlyValGlyValArgTyrAlaSerPro-----ValGly 892
Db 9428 GATCCCGACCAAGCGCCAAATGCCAGGTGTTTACTCTCTCCCGGATTCACGCCAAGGC 9487
Qy 893 GlnValArgValAspValAlaThrGlyValLysGluGluGlyAsnProIleLysLeu 911
Db 9488 CACTTTTCCATCGAGCGCCACCGGGGTGATCCGCTGGAAAGCCGCTGCAGGTC 9544
RESULT 15
US-09-815-242-8512
; Sequence 8512, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
```

```
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8512
; LENGTH: 17388
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(17388)
US-09-815-242-8512
Alignment Scores:
Pred. No.: 0.0309 Length: 17388
Score: 133.00 Matches: 165
Percent Similarity: 31.24% Conservative: 144
Best Local Similarity: 16.68% Mismatches: 348
Query Match: 2.81% Indels: 332
DB: 10 Gaps: 45
US-09-914-168-2 (1-919) x US-09-815-242-8512 (1-17388)
Qy 51 GlnAlaLysAlaGlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaArg--- 69
Db 2374 CGGTAAACCAACATACACCAACAAATTAATTTCTGGAATACCGTTCATCACCATTGGTTAT 2433
Qy 70 -----LeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeu 84
Db 2434 TTACAAAAAGTACCTGTTAATATTTCTGGAATACCGTTCATCACCATTGGTTAT 2493
Qy 85 AspValValAsnPheAspAspGlnSerProIleSerArgIle-----GlyGluGln 101
Db 2494 TTAGTT-----CGGCCAACAAATCAAACTACGAATGGTGTAGTGAG 2535
Qy 102 SerProProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGlu 121
Db 2536 GCAGATCAAAATACCATCTGGTTATACGATACCTGGACTGGTACCTGATGGGTCAT 2595
Qy 122 GluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyr----- 138
Db 2596 AACACAATTACT-----ATACACCAAAAGATTACGTAGTATT 2634
Qy 139 IleProGluTyrGlnGlyGluGln----- 146
Db 2635 ATACCACCAAGTA---GGTAAACAAATTAGAGCAGTAGTTTATTATAATAAAGTAGTTGCA 2691
Qy 147 -----ProAsnSerGluValValValProProThrLeuGluProGluLysProGly 163
Db 2692 TCTAATATGAGTAATGCTGTTACTATTTTGGCCAGATGACATTCACCAACA----- 2742
```

QY 164 LeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeu 183
Db 2743 -----ATCAATAATCCTGTTGGAATA 2763
QY 184 LysAlaLysPheTyrGlnSerSerGlnSerGlyThrSerAlaIleGlySerSerHis 203
Db 2764 AATGCCAAATACTATCAGGCGGAGAGTCAACTTTACAATGGGTCTCTGTAGATACAT 2823
QY 204 Gln-----LysThrGluProTyrAlaAsnIleLysAlaAla---LeuGluAspIleThr 220
Db 2824 TCTGGATTAAAGTACAACTATTACGACATGCCAAGTGGTGGACATCAAAATTTAACT 2883
QY 221 GlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuVal 240
Db 2884 AATATCCGAC-----AACAAAAACGGCTCATTTAGCTATTACTGCTAGAGTGAGTATGAAT 2937
QY 241 AlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGly 260
Db 2938 CAGGCA-----TTTAACAGTGATATTAAATTTAAAGTGTGAGGACAGAT 2982
QY 261 GluValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaVal 280
Db 2983 AATCTCAATAATACGACAAATGAT---AGTCAATCTAAACATGTGTCAATTCTAGT 3039
QY 281 GluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValPro 300
Db 3040 AAAATTAGT-----GAAGATGCTCATCCG 3063
QY 301 LeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGlu 320
Db 3064 ATTGTATTAGAAAT-----ACTGAGAAAGTTGTAGTAGTC 3099
QY 321 AsnAlaSerAlaGlu----- 325
Db 3100 AATCCGACTCTGTATCTAATGATGAAACGAAAGATAATTACTGCCTTTATGAATAAA 3159
QY 326 -----HisGlyTyrPheAspGlyArgTrpLeuAspArgSerValAspValIle 341
Db 3160 AACCAAAATATAAGAGGATATTTA-----GCATCAACTGATCCAGTA 3201
QY 342 LeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPhe 361
Db 3202 ACAGTCGATAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3261
QY 362 AspGluValValPhePheThrIleAspPro-----LysThrAsnGlnLeuThrThrAsp 379
Db 3262 GATGCTACAAATCTGATGACATACGAACCAAGTGTGAATCTGAATATCAAACTGCCAAT 3321
QY 380 ProAspLysLeuProValLysArgGluLeuGluGlnLeuThrValAsnMetGly 399
Db 3322 GCTGTAAACACGCAAG-----GTAACGATTGCTAAAGGA 3357
QY 400 GluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyr 419
Db 3358 CAATCATTTAATATGATGATATATAA-----CATAT 3390
QY 420 PheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGln 439
Db 3391 TTTTACTTAAGTAATGGACAAGCTATT-----CCAAAGT 3423
QY 440 ValSerPheGluGlnSerSerSerArgThrGluPro---AlaGlnValAspGluSer 458
Db 3424 GGAAACATTTACAAATATTACATCTGATAGAACTATTCCAACTGCACAA----- 3471
QY 459 ThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSer 478
Db 3471 ----- 3471
QY 479 ProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaLaLys 498
Db 3472 -----GAAGTTAGTCAAAATGAATGAGGTACAGCTATATCATATAGTTGCTTCAAT 3525

QY 499 AlaArgHisLeuTyrAspMetProAspAsp-----ArgValLeuAlaIleAsn 514
Db 3526 GCATATCAT---AAAGCACTGAAGATTTCATATTAGTTTAAAAAATCGTTGATGTGAAA 3582
QY 515 HisAspAspGlyValAsnArg-----SerIleLeuGlyArgIleSerAspAla 530
Db 3583 CAACCTGAAGCGCATCAACGTTGATATCGTACGTCAACATATGATTTAACTACTGATGAA 3642
QY 531 ValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeu 550
Db 3643 ATCTCAAAAGTAAACAAGCTTTTATT---AATGCCAAATAGAGATGTAATTTACTCTT 3696
QY 551 ProGlu---ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr----- 566
Db 3697 GCCAAGGTGATATTTCAGTTACAATAACACCTAATGCTGCTAATGTAAGTACTATTACA 3756
QY 566 ----- 566
Db 3757 GTAATAATTAAAGGTGCGATTAAACGAATTCATTCGCGTCTAACCTAGCTAATATCAAT 3816
QY 567 -----GlnSerLysLysValProLeuTyrValPheValAla 578
Db 3817 TTCTTGGCTGGGTTAATTTCCCAAGATTATACAGTGACATGACGAATGTAANAAT 3876
QY 579 SerAspLysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAspThrGlyThrArg 598
Db 3877 GCAACAGACCAACAGATGGT-----GGTTTATCATGTCGTCGATGAC----- 3918
QY 599 LeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlu 618
Db 3919 -----CATAAATCTTTATTTATCGTTATGCTATGCTATGCTAGTCAGGCACACA 3963
QY 619 LeuArgLeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHis 638
Db 3964 ATTCAACTAATGATATTTTAAACGATGCTAAACCGCATCTACTACAGTGCCT----- 4014
QY 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGly----- 656
Db 4015 GGTTCGCTAATAATATTACTGTAATGAAAGACGACAGCAGAGGTGGGAAGACCA 4074
QY 657 ---HisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArg 675
Db 4075 AACTATAGACAACACTGGTTATTTCACAAGCAATGCATCATCTGATGCCAA----- 4125
QY 676 SerIleIleGlnAsnGlyGlyTrpAsnArgThrThrSerLeuArgTyrArgLeuAspLys 695
Db 4126 -----AGCAATATACGTTG----- 4140
QY 696 LeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheValAsnGly 715
Db 4141 -----AATGGT 4146
QY 716 LysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsn 735
Db 4147 CAAGTGATTCAA-----ATATTAGAC--- 4167
QY 736 LeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGly 755
Db 4168 ATCATCAACCTTCAACCGCTTATGTTGGGCAACCTGTTACAAATTTCAAACTACTCGTCCA 4227
QY 756 LeuValSerAspAla-----AsnMetAlaIleAlaArgAlaGlyIleSerGlyVal 772
Db 4228 AACCATAGTAAACCACTGTTTAACTAACGACCAACCGCAGCTAATGTTGCTGGCGCA 4287
QY 773 TyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyLe 792
Db 4288 TTTTACAATT---CACCACGTTGTAAGAAAGTAATTTCTACACATAATGCAAGTGATGCACTT 4344
QY 793 GlnAlaGlyTyrIleTrpSerAspAsnPheAsnHisValProTyr-----ArgLeuArg 810
Db 4345 TATAAAGCACAGTTATCTTAAACGCCCATATGGGCCCAAAACAATATGTTGAACATTTAAAT 4404
QY 811 PhePheAlaGlyGlyAspGln-----SerIleArgGlyTyrAlaHisAspSerLeu 827


```
Db 691 GCCCTCCGTGAC-----ATGCTCAAAACACACACTCTA----- 723
Qy 467 ValGluLeuThrAspGlyLeuMetAspIleSerProIleGluPheSerAlaSerAsn 486
Db 724 ---CAACTCTTTCCCTCCTGCCATGAC---AAACCGCAAGCTTCACAAAGACGAG 777
Qy 487 LeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetPro 506
Db 778 ATTCTGTGCGAAAGATTAAAGTCTTT-----AAAACTCTATCAT---CCA 822
Qy 507 AspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArg 526
Db 823 ACTGATGAAGAACTCAAGAACTTTATCCGTGGCAATACCCCTCT-----GGTAAG 876
Qy 527 IleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu 546
Db 877 ATT---GATGCGATCAATACATCTCTATCTGTCAGCGCAAT----- 918
Qy 547 ValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr 566
Db 919 -----GTGAATCCAGAATCAACAACACTGAAACCTTT 948
Qy 567 GlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGln 586
Db 949 ACATCTGGTGCC-----TTCTTTGTAGACAGCGATCGATTCCGT----- 987
Qy 587 IleGlyLeuGlyTyrPglySerAspThrGlyThrArgLeuValThrLysPheGluHis--- 605
Db 988 ---GGTCTCTCTTTCTTTTCCGTACAGGTAAACGACTGACTGAAAAAGGAACATCATGTC 1044
Qy 606 AsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaLeuLeuArgLeuSerGluAspLys 625
Db 1045 AACATCGTC----- 1053
Qy 626 LysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArg 645
Db 1054 -----TTTAAACAAATGGATTCTATCTTTGGAGAACCACTTGCT 1092
Qy 646 AlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSer 665
Db 1093 CCAATATTT-----TTGACCATCTATATTCAACCAACAGAGGCTTCTCTTAGC 1143
Qy 666 ThrArgThrLeuGluHisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArg 685
Db 1144 CTAAATGGGAAGCAAGTAGGAGAAGATTAACTTGGCTCCTAAC----- 1188
Qy 686 ThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrp 705
Db 1189 -----TCACTTGATTACCGTACAGATCGCACTGCAACTGGTCTCTCCAGAACCATAC 1242
Qy 706 GlnAspLeuProValAspPheValAsn 714
Db 1243 GAAATTTGATTTATGATGCTCTAAAT 1269

RESULT 17
US-09-815-242-7037
; Sequence 7037, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Cart, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELTRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
```

```
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7037
; LENGTH: 2571
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2571)
; US-09-815-242-7037

Alignment Scores:
Pred. No.: 0.00211 Length: 2571
Score: 132.50 Matches: 109
Percent Similarity: 36.35% Conservative: 96
Best Local Similarity: 19.33% Mismatches: 212
Query Match: 2.80% Indels: 147
DB: 10 Gaps: 25

US-09-914-168-2 (1-919) x US-09-815-242-7037 (1-2571)
Qy 28 AlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisAspThr 47
Db 343 GCTTTAGAAGAAGAGGAAACGATCAAGCATATTTTGAAGAACTGGCGTGAAGAA 402
Qy 48 AlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGluGlnIleGln 67
Db 403 CAAATTTCCCAAGCT-----ATTACGACCATTTAGA 432
Qy 68 AlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspVal 87
Db 433 GGGGACAAAACGCTGAACGATCAAAATGCAGAAAGAAAGCAGACAAAGCGCTTGAA 492
Qy 88 AsnPheAspAspGlnSerProIleSerArgIleGlyGluGlnSerProProLeuGlyLeu 107
Db 493 ACGATT---GATTTAACCCCTCGTCGAGAAAGTGGCAAACTTGATCCTGTATTGGCG 549
Qy 108 AspMetSerValIleGluGluThrThrProLeuSerLeuGluLeuPheAlaGlnGlu 127
Db 550 GAT-----GAAGAAATTT----- 561
Qy 128 SerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnPro 147
Db 562 -----CGTCGAGCCATTCAGTATTACAAACGTCGTACCAAA 597
Qy 148 AsnSerGluValValProProThrLeuGluPro-----GluLysProGlyLeuIle 165
Db 598 AATAACCTGTGTTAATTGGT-----GAACGAGGTAGGAGGAAACCGCGATTTGTC 648
Qy 166 LysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArg---LeuLys 184
Db 649 GAAGGCTTGGCACAGCGCATCGTAAAC-----GGCGAAGTGGCAGAGGTTTCAAA 699
Qy 185 AlaLysPheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGln 204
Db 700 AATAACCGTGTCTTTCATTAGATATGGGG-----CGCTTGATTGCTGTCGCAAAATAT 753
Qy 205 LysThrGluProTyrAlaAsnIleLysAlaLeuGluAspIleThrGlnGlnSerAla 224
Db 754 CGTGTGTAATTTGAAGAACGTTTAAAGCAGCTACTCAATGAACATTTTCAAGAAAGAAAGGT 813
```



```
Qy 225 MetAspLeuAnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAla 244
Db 814 CGCGTTATCTCTC---TTTATTGACGAAAT---CATACTATGTCGCGCGGGTAAAC 867
Qy 245 ValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspVal 264
Db 868 GATGGTCGGATGGATGCGGGT---AATTGTGTTAAACCAAGTTTGGCAGGACGCAA--- 921
Qy 265 IleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGly 284
Db 922 ---TTACATTCGCGGGTGCACACTTATAGATCAATATCGTCAATAT---ATCGAAAA 975
Qy 285 GluGlyAlaAspAspLysAlaPhe-----ThrThrValAlaAsp 297
Db 976 GATGCCGCACCTGAACCCGCTTCCAAAAAGTCTTGTGGACGAACCAAGTGTAGAAGAT 1035
Qy 298 GluValProLeuLeuIleGly-----AspValPheHisHisGlyLysTyr 312
Db 1036 ACCATTCCGATCTACGTGGTTTGAAGACGGTTATGAAATTCATCATCAGTGGATATT 1095
Qy 313 GluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArg 332
Db 1096 ---ACTGACCCAGCAATGTGCGTCGACCAACGCTTTCACATCGTTATATTTCGATCGT 1152
Qy 333 TrpLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIle 352
Db 1153 CAG-----TTACCAGATAAAGCCATTGATTTC----- 1179
Qy 353 TyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIleAspProLys 372
Db 1180 -----ATCGATGAAGCA 1191
Qy 373 ThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGluGln 392
Db 1192 GCGCTAGCATTCGTATGGAATAGATTCTAAACCTCAACCGCTTGATCGCTTTGAACGT 1251
Qy 393 LeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsn 412
Db 1252 CGTATTATCCAAATTAATAATTGGAACAACAACGCTGTACAAAA-----GAAGAA 1299
Qy 413 AspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArg 432
Db 1300 GACGAACAGTCGCAACGCTTTAGAAATGTTAGAGAAGAAATTCGCTGAAAAAGACGT 1359
Qy 433 GluGlnIleGlnAsnAspGlnVal-----ThrLeuGluProValIle 464
Db 1480 GATTTACGGAATGCTGATGTCATATGCGCGCATCCCTGATCTTGAAAGCAACTT 1539
Qy 465 GluThrValGluLeuThrAspGlyIleLeuMet----- 475
Db 1540 GAGCAACTGAAACCCGGAAGGAAAAGAAATGACGGTTTACGCTATCCGTCACAGAT 1599
Qy 476 ---AspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeu 494
Db 1600 GAAGAAATCGCAGAGTGCTTCTTAAAGCCACAGGCAATTCCTGTATCAAAAATG----- 1653
Qy 495 ValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeuAlaIleAsn 514
Db 1654 -----ATGGAAGCCGCAAGAAAAAATCTTCGCT 1693
Qy 515 HisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaVal 534
Db 1684 ATGGAAGATGACTACATAACGAGTGATTTGGTCAA---GAAGAAGCGGTTGATCGCGTA 1740
```

```
Qy 535 AlaArgAlaIle 538
Db 1741 GCAAACGCGATT 1752

RESULT 18
US-09-815-242-4761
; Sequence 4761, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsep, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4761
; LENGTH: 7434
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4761

Alignment Scores:
Pred. No.: 0.0112 Length: 7434
Score: 132.00 Matches: 176
Percent Similarity: 32.22% Conservative: 123
Best Local Similarity: 18.97% Mismatches: 326
Query Match: 2.79% Indels: 304
DB: 10 Gaps: 40

US-09-914-168-2 (1-919) x US-09-815-242-4761 (1-7434)

Qy 2 SerLysProValLeuPheAlaAsnArgSerPheMetProValAlaLeuAlaAlaTyrLeu 21
Db 1423 ACTAAGAAGTTTATTTCACAGATAATCT-----TTA 1455
Qy 22 ProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIle----- 38
Db 1456 AAATTATCATATAAAGTTAATGTTTGGCAATATCGATACACCTAAATAATTTGATTTAAT 1515
Qy 39 -----IleAsnHisValProAlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyAsn 56
Db 1516 GAAAAATTAACATATCTACTGCTTCAGATGTTCTTAATAATATGCGCAA----- 1566
Qy 57 ProProValLeuLeuThr-----ProGluGln 65
Db 1567 CCAGAAGTTTACACTTAACCTGCAGATCCATTTTCAGTAGCGGTTGAATGAACAAGATCGG 1626
Qy 66 IleGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAla----- 77
Db 1627 TTGCAACAACAAGTAAACTCACAAGTTTGATAATAGTATTACACAACAGCATCAATTGCA 1686
```

Qy	78	-----LysProGlnSerGlnAlaLeuAspValValAsnPheAsp-	-----	90
Db	1687	GAATACAATAAACTTAAACAACAGCAGATACTATTTTAAATGAAGATGCGAATCATGTT	: : : : :	1746
Qy	91	-----AspGlnSerProIle-	:	95
Db	1747	AAAACTGCAATCGTCATCTCAAGCGGATATTGATGGTTAGTAACATAAATTTACAAGCT	:	1806
Qy	96	SerArgIleGlyGluGlnSerProProLeuGlyLeuAspMetSerValIleGluGluThr	: : : : :	115
Db	1807	GCATTAAATTCATAATCAAGCAGCAATTCGTGAATTAGATACTAAGCTCAAGAAAGGTT	: : : : :	1866
Qy	116	ThrProLeuSerLeuGluGluLeuPheAlaGlnGluSer-	-----ThrGlu	130
Db	1867	ACAGCAGCACAAAGATAAAAAGTTACGCAAGATGAAGTTGCAGCAGCTTGTAACCTAAA	: : : : :	1926
Qy	131	MetGlyIleAsnProAsnAspTyrIleProGluTyr-	-----GlnGly	144
Db	1927	ATTAAACAATGATAAAATAATGCATCCGAGAAATTAATAACAACAATCAGACACAGGT	: : : : :	1986
Qy	145	-----GluGlnProAsn-	-----SerGluValValValProThr	156
Db	1987	GTCAAACTGAAAAAGATAATGATCGCAGTGTAGAACAAAGATGATTACACCAACA	: : : : :	2046
Qy	157	LeuGluProGlu---LysProGlyLeuIleLysArgLeuTyrAlaArg-	-----	171
Db	2047	GTTAAACCTCAAGCGAAACAAGATATTATCCAAAGCAGTTACAACCTCGTAACAACAAT	: : : : :	2106
Qy	172	-----LeuPheAsnAspGlyValAsnLys	-----	179
Db	2107	AAAAAGTCAATGCATCAATACAAGATGAAAAGATGTAGCAAAATGATAAAATGGTTAA	: : : : :	2166
Qy	180	ValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGly-	-----	194
Db	2167	ATTGAAACAAAGGCAATTTAAAGATATTGATGCAGCAACAACAATGCACAAGTAGAAGC	: : : : :	2226
Qy	195	---GluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLys	-----	213
Db	2227	ATTHAACAAACAAATCAATGATATTATCAACTACACCTGCTCAACAGCTTAAGCA	: : : : :	2286
Qy	214	AlaAlaLeuGlu-	-----	217
Db	2287	GCAGCTCTTGAAGAAATTTGACGAAGTTGTCACGACAAAATTTGATCAAGCACCTTTAAAT	: : : : :	2346
Qy	218	---AspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGln	-----	236
Db	2347	CCTGATACAAACAATGAAGAAGTAGCGGAA	: : : : :	2397
Qy	237	ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIle-	-----AspLeuSerIle	254
Db	2398	GCTAAAGTTTCTGGTGTAAAGCAATTGACGCACACGACTGCACAAGATTTTAGAAAGA	: : : : :	2457
Qy	255	IleArgAsn-	-----	257
Db	2458	GTTAAACACGAACAAATCTCAAAAATTTGAAAATATTACTGACTCTACGCAACAAAAATG	: : : : :	2517
Qy	257	-----	-----	257
Db	2518	GATGCTTATATGAAGTTAAACAAGCTGCAACAGCTAGAAAAGCTCAAAAATGCTACAGTT	: : : : :	2577
Qy	258	-----SerIleGlyGluValAspValIleIleHisAspLeuGlyGlu	-----	271
Db	2578	TCAATGCCAANAATCAAGAGTAGCAGACGCTGATGCAGCAGTAGATGCAGCTCAAAAG	: : : : :	2637
Qy	272	ProValTyrIleAspTyrArgAlaValGluValArgGlyGlyGlyAlaAspAspLys---	-----	290
Db	2638	CAAGGTTTACATGCATCCAAGTTGTTAAATCAAACAGGAAGTTGCTGATACAAATCA	: : : : :	2697
Qy	291	-----AlaPheThrThrValAlaAspGluValProLeuLeuIle	-----	303
Db	2698	AAAGTATTAGATAAAATCAATGCAATTCAACACCAAGCAAGGTTAAACCT-----GCA	: : : : :	2751

Qy	304	GlyAspValPheHisHisGlyLysTyThrGluThrLysLysAsnLeuIleGluAsnAlaSer	329
Db	2752	GCTGCTACGGGAAGTAGAAAAACGCGATATATAATACGTAACAAGAATAATCAAAATAGCAAT	2811
Qy	324	AlaGlu-----HisGlyTyThrPheAspGlyArgTrpLeuAsp	335
Db	2812	CCTTCAACTACAGAAGAAAAACAAAGCTGCATATACAGAAATTAGATTACTATAAACACGACGAA	2871
Qy	336	ArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyTyrAsn	354
Db	2872	GCAAGAACAAATCTTGATGCTGCAAAATACAAACAGTGATGTAACACACAGCTAAAGACAAAT	2931
Qy	355	-----ThrGlyThrGlnTyArgPheAsp-----	362
Db	2932	AGTATTGCTGCAATTATCAAGTTCAGCTGCCACAACCTAAGAAAATCGGATGCAAAAACGC	2991
Qy	363	GluValValPhePheThrIleAspProLysThr-----AsnGlnLeuThr	377
Db	2992	GAAATCGCTCAAAAAGCAAGTGAACGTAATAACAGCAATTGAAGCAATCAATGATTCGACT	3051
Qy	378	ThrAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuThrValAsn	397
Db	3052	ACTGAAGAACAAACAA-----GCAGCGAAGACAAAGTGGATCAAGCAGTAGTACTGCCA	3105
Qy	398	MetGlyGluAlaTyArgAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThr	417
Db	3106	AACGCTGATATAGATAAT-----GCTGCAGACAAACAATGATGTGGATTAATGCCA	3153
Qy	418	ArgTyThrPheAsnMetValAsnThrGluIleValPheProGlu-----	431
Db	3154	AAAACTACAAATGAAGCTACAAATCGCAGCGCATTTACACCTGATGCAAAATTTTAAACACGACGA	3213
Qy	432	-----ArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSer	446
Db	3214	GCAAAACACGCAATTCGAGATAAAGTACAAAGCTCAAGAAACAGCAANTTGATGGAAATAAC	3273
Qy	447	SerSerArgThrGlu-----ProAlaGlnValAspGluSerThr	459
Db	3274	GGCTCAACAACCTGAAGAAAAAGCAGCTGTAAACAAACAAGTTCAACCTGAAAAAACAAACA	3333
Qy	460	LeuGluProValIleGlu-----ThrValGlnLeuThrAspGlyIleLeuMet	475
Db	3334	GCTGATCGCCGCAATAGATGCAGCACATACAAATCGGGAAGTTGAAGCGGCTAAAAAAGCA	3393
Qy	476	AspIleSerProIleGluPhe-----SerAlaSerAsn	486
Db	3394	GCAATTCGTAATAATTTGAAGCGATTACGCGACCAACAACAACCTAAAGATAAATGCGAAAGAA	3453
Qy	487	LeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyArgPheMetPro	506
Db	3454	GCAATTCGTACGAAACGGAATGAACCTAAACAGCAATCGCTCAACCGCAAGACATTACT	3513
Qy	507	AspArgValLeuAlaIleHisHisAspGlyValAsnArgSerIleLeuGlyArg	526
Db	3514	GCTCAAGAAAATTCAGCGCGCTAATGCGGAC-----	3543
Qy	527	IleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu	546
Db	3544	GTAGATAATGCTGTGACACAAAGCAAAATAGCACATTTGAAGCTGTCTAAATGATCAAAATGAT	3603
Qy	547	Val-----IleAsp-----	549
Db	3604	GTACACCAACGGAACGACAGGTGAAATAGTATTGATCAAGTAACACCAACACGTTAAT	3663
Qy	550	-----LeuProGluArgThrAlaLeuAlaAsnArgLys-----	560
Db	3664	AAAAAAGCAACTGCAGGTAAATGAAATCACAGCAATTTTAAATACAAATTCGAAGAGATT	3723
Qy	561	-----ThrProAlaAspValTyThrGlnSerLysLysValProLeuTyValIlePheValAla	578
Db	3724	CAAGCTACGCCGATGCACAGATGAAGAAAAAACAAGCAGCT-----GATGCT	3771
Qy	579	SerAspLysProArgAspGlyGlnIleGlyLeuGlyTyThrGlySerAspThrGlyThrArg	598

Db 2398 GCTAAAGTTTCTGTTAAAGCAATTGAACGGACACGACTGCACAAAGATTTAGAAAGA 2457
Qy 255 IleArgAsn----- 257
Db 2458 GTTAAACGAAGAAATCTCAAAAATTGAAATATTACTGACTCTACGCAACAAAATG 2517
Qy 257 ----- 257
Db 2518 GATGCCCTAATGAAGTTAAACAAGCTGCAACAGCTAGAAAAGCTCAAAATGCTACAGTT 2577
Qy 258 -----SerIleGlyValValAspValIleIleHisAspLeuGlyGlu 271
Db 2578 TCAATGCAACAATGAAGAGTAGCAGAAGCTGATGCCAGCAGTAGATGCGAGCTCAAAAG 2637
Qy 272 ProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAspLys--- 290
Db 2638 CAAGTTTACATGACATCCCAAGTTGTTAAATCAAAACAGCAAGTTGCTGATACAAAATCA 2697
Qy 291 -----AlaPheThrThrValAlaAspGluValProLeuLeuIle 303
Db 2698 AAGATTATGATAAAATCAATGCAATTCAACACACAGCAAAAGTTAAACCT-----GCA 2751
Qy 304 GlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSer 323
Db 2752 GCTGATACGGAAGTAGAAAACGATATATACACGTAAACAGAAATTCAAAATAGCAAT 2811
Qy 324 AlaGlu-----HisGlyTyrPheAspGlyArgTyrLeuAsp 335
Db 2812 GCTTCAACTACAGAAGAAAACAGCTGCATATACAGAAATAGATATAAAAACGAAGAA 2871
Qy 336 ArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAsp--- 354
Db 2872 GCAGAAACAATCTGATGCTGCAAAATACAAACAGTATGATACACAGCAGCTAAAGACAAT 2931
Qy 355 -----ThrGlyThrGlnTyrArgPheAsp--- 362
Db 2932 AGTATTGCTCAATTAAATCAAGTTCAAGCTGCCACACTAAGAAATCGGATGCAAAAGCG 2991
Qy 363 GluValValPhePheThrIleAspProLysThr-----AsnGlnLeuThr 377
Db 2992 GAAATGCTCAAAAAGCAAGTGAACGTAAACAGCAATTTGAAGCAATGAATGATTCGACT 3051
Qy 378 ThrAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrValAsn 397
Db 3052 ACTGAAGACAACAA-----GCAGCGAAAGACAAGAGTGTGATCAAGCAGTAGTTACTGCA 3105
Qy 398 MetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThr 417
Db 3106 AACGCTGATATAGATAT-----GCTGCACCAACAATGATGTGGATAATGCA 3153
Qy 418 ArgTyrPheAsnMetValAsnThrGluIleValPheProGlu----- 431
Db 3154 AAAACTACAATGAAGCTACAACTCGCAGCCCTACACCTGATGCAAAATGTTAAACCCAGCA 3213
Qy 432 -----ArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSer 446
Db 3214 GCAAAACAACAATGTCAGATAAGTACAAAGCTCAAGAAACACCAATTTGATGGAAATAAC 3273
Qy 447 SerSerArgThrGlu-----ProAlaGlnValAspGluSerThr 459
Db 3274 GGCTCAACAACCTGAGAAAAGCAGCTGCTAAACAACAAGTTCAACTGAAAAACAACA 3333
Qy 460 LeuGluProValIleGlu-----ThrValGluLeuThrAspGlyIleLeuMet 475
Db 3334 GCTGATGCCCAATAGATGACGACATACAAATGCGGAAGTTCAAGCGGCTAAAAAAGCA 3393
Qy 476 AspIleSerProIleGluPhe-----SerAlaSerAsn 486
Db 3394 GCAATGCTAAAATGAAGCGATTTCAGCCGACCAACAACACTAAAGATAATGCGAAAGAA 3453
Qy 487 LeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetPro 506
Db 3454 GCAATTGCTACGAAGCAAGATGAACGTAAACACAGCAATCTGCTCAACGCGAAGACATTACT 3513

Qy 507 AspAspArgValLeuAlaIleAsnHisAspAspGlyValAlaAsnArgSerIleLeuGlyArg 526
Db 3514 GCTGAAGAAATTCAGCGGCTAATGCGGAC----- 3543
Qy 527 IleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu 546
Db 3544 GTAGATAATGCTGTGACACAAGCAAAATAGCAACATTTGAAGCTGCTAATAGTCAAAATGAT 3603
Qy 547 Val-----IleAsp----- 549
Db 3604 GTAGACCAACGAAACAGCAGAGTGAANAATAGTATTGATCAAGTAACACCAACAGTTAAT 3663
Qy 550 -----LeuProGluArgThrAlaLeuAlaAsnArgLys----- 560
Db 3664 AAAAAAGCAACTGCACGTAAATGAATCACACGAATTTTAAATTAACAAATTTGCAAGAGATT 3723
Qy 561 -----ThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAla 578
Db 3724 CAAGCTACGCCAGATGCACAGATGAAGAAAACAAACAGCAGCT-----GATGCT 3771
Qy 579 SerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArg 598
Db 3772 GAAGCAAAATCTGAAATGCTAAGCAAAATCAAGCCATTTTCAGCAGCACTACTAAGCGCA 3831
Qy 599 LeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlu 618
Db 3832 CAAGTT-----GATGAAGCTAAAGCAAAATGCAGAA 3861
Qy 619 LeuArgLeuSerGlu---AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSer 637
Db 3862 GCAGCGATTAAATCGGTACACCAAAAGTTGTGAAGAAACAAAGCGGCTAAAGATGAAATT 3921
Qy 638 HisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPheGlyHis 657
Db 3922 -----GATCAATTTACAGCAACG-----CAAACAATGTTATCAATAAT 3960
Qy 658 SerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSerIle 677
Db 3961 -----GATCAGAAGCTGTACAAACAGAAAGAAAGCAAGC-AGCTATTCAACA 4004
Qy 678 IleGlnAsnGlyGlyTyrAsnArg 685
Db 4005 ATTAGCAACAGCAGTTTACAGACGC 4028
RESULT 20
US-09-815-242-4648
; Sequence 4648, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27

```
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4648
; LENGTH: 6077
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4648
```

```
Alignment Scores:
Pred. No.: 0.00938 Length: 6077
Score: 131.50 Matches: 184
Percent Similarity: 33.76% Conservative: 135
Best Local Similarity: 19.47% Mismatches: 345
Query Match: 2.78% Indels: 281
Db: 10 Gaps: 48
```

US-09-914-168-2 (1-919) x US-09-815-242-4648 (1-6077)

```
Qy 48 AlaIleAsnGlnAlaLys-----AlaGlyAsnProProValLeuLeuThrProGlu 64
Db 942 GCAGTGAATGACGCTAAACAAACAACTTACATGGTGTGATCAAAATTTAGCTCAAGATAAGCAA 1001

Qy 65 GlnIleGlnAlaAargLeu---AsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAla 83
Db 1002 CGTGCACAGAAACGTTAAATAACTTGTCTAACTTGAATACA---CCACAGCTCAAGCA 1058

Qy 84 LeuAsp-----ValValAsnPheAspAspGlnSerProIleSerArgIleGlyGluGln 101
Db 1059 CTTGAAATCAAAATCAATTAATGCACAACTCGTGGCGAAGTAGCACAAATAATTAAGTGG 1118

Qy 102 SerProProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGlu 121
Db 1119 GCACAAAGCATTAAACCAAGCAATGAAGCATTTACGTAAT-----AGCAITCAA 1166

Qy 122 GluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGlu 141
Db 1167 GAT-----CAACAGCAACAGACAGCGGT-----AGCAAGTTTATCAATGAA 1208

Qy 142 -----TyrGlnGlyGluGlnProAsnSerGluValValValPro 154
Db 1209 GATAAGCCACAAAAAGATGCTATACCAGCAGCAGTACAAACATGCCAAAGATTGATTAAT 1268

Qy 155 ProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsn 174
Db 1269 CAACTAGCAATCCA-----ACGCTTGATAAGCACAAAGTTGCAAAATTGACA 1316

Qy 175 AspGlyValAsnLysValPro-----ArgLeuLysAlaLysPheTyrGlnSerSer 191
Db 1317 CAAGGTGTTAACCAAGCTAAAGATAACCTACATGGTGTGATCAAAACCTTCCAGACGATAA 1376

Qy 192 GlnSerGlyGluThr-----SerAlaIleGlySerSerHisGlnLys 205
Db 1377 CAACATGCGAGTTACTGATTAAATCAATTAATAGTTTAAATAATCCGCAAGCTCAAGCT 1436

Qy 206 ThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGlu----- 222
Db 1437 CTTGAAAGTCAATTAACACCCGCAACTCGTGTGATGAAGTAGCGCAAAATTAAGTCTGAA 1496

Qy 223 ---SerAlaMetAsp-----LeuAsnGlySerIleProArgLeuArgGln 236
Db 1497 GCACAAGCGCTTGATCAAGCAATGCAAGCATTACGTAAATAGTATTCAAGATCAACAAACA 1556

Qy 237 ThrAlaLeuValAlaAlaAargAlaValGlyTyrTyrAspIleAspLeuSerIleIleArg 256
Db 1557 ACAGAA---TCTGGTAGCAAGTTTATCAATGAAGATAAGCCACAAAAAGATGCATACCAA 1613

Qy 257 AsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAsp 276
Db 1614 GCAGCAGGTACAAACATGCCAAAGATTTAATTAACCAACAGGTAATCCAACA---CTCGAT 1670
```

```
Qy 277 TyrArgAlaValGluValargGlyGluGlyAlaAspAspLysAlaPheThrThrValAla 296
Db 1671 AAATCACAGTTGAACAATTA-----ACACAGCAGTACAACTGCAAAA 1715

Qy 297 AspGluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLys 316
Db 1716 GATAACCTA-----CATGGTGATCAAAAACTTGTCTGT 1748

Qy 317 AsnLeuIleGlu-----AsnAlaSerAlaGluHisGlyTyrPheAspGly 331
Db 1749 GATCAACCAACAGCAGTAAAGCTGTTAATGCATTACCAAACTTAAATCATCGGCAACAA 1808

Qy 332 ArgTrpLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeu 351
Db 1809 CAAGCTTTAACTGATGCTATAAATGCA---CGCCCTACAGAACACA---GAGTTGCACAA 1862

Qy 352 IleTyrAspThrGlyThrGlnTyr-----ArgPheAsp 362
Db 1863 CATGTTCAAACTGCTACTGAACCTTGATCAGCGATGGAACACATTGAAAAATAAAGTTGAT 1922

Qy 363 GluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLys 382
Db 1923 CAAGTGAATACAGATAAGCTCAACCAATTAACACTGAAGCGTCAACTGAT----- 1973

Qy 383 LeuProValLysArgGluLeuLeuGluGlnLeuLeuThrVal----- 396
Db 1974 -----AAAAAGAGCAGCTAGATCAAGCGTTTACAAGCTGCAGAAAGCATTACAGAT 2024

Qy 397 ---AsnMetGlyGluAlaTyrAsnLeuGlnAlaVal---ArgAlaLeuSerAsnAspLeu 414
Db 2025 CCAACTAATGGTTCAAAATGCCGAATAAAGACGCTGTAGAACACGATTAACCTAAG----- 2078

Qy 415 IleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGln 434
Db 2079 CTTCAAGAAAAAGTAAATGAGTTAAATGGTAATGAGAGAGTGCCTGAAGCTAAAGCACAA 2138

Qy 435 IleGlnAsn-----AspGlnValSerPheGluGlnSerSerSerSerArgThrGluPro 452
Db 2139 CGGAAACAAACTATTGTACCAATTTAGCACATTATAATGCTGTGATCAAAATTCGAACCTGCTAAA 2198

Qy 453 AlaGlnValAspGluSerThr---LeuGluProValIleGluThrValGluLeuThrAsp 471
Db 2199 CAAAACTTTGATCAAGCAGCAAACTTCAACCAATTCGCTGAATAGTAGAT----- 2249

Qy 472 GlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLys 491
Db 2249 ----- 2249

Qy 492 LeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeu 511
Db 2250 -----CAAGCAACGCAATTGAATCAATCAATCTATGATCAATTCGAACAA 2291

Qy 512 AlaIleAsnHisAspAspGlyValAsnArgSerIle-----LeuGlyArgIleSerAsp 529
Db 2292 CGAGTCAATGAACATGCTAAACGTTGAGCAAACTAGATATTACACGCAAGCAGATTCAGAT 2351

Qy 530 AlaValSerAlaValAlaAargAlaIleLeuProAspGluSerGluAsn----- 545
Db 2352 AAGCAAAATGCTTATAACACAGCTATTGCA-----GAAGCTGAAAATGATTATAAACAA 2405

Qy 546 -----GluVal 547
Db 2406 AATTCGAATAAGCAACAGTGGATCAAGCAGCTTCAAAACATTTTAAATGCAAAACAGCA 2465

Qy 548 IleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGln 567
Db 2466 TTAATGCTGATGAACGCTGTAGCACTTGTCTAAAAACAAATGTTAAACATGACATCGACCAA 2525

Qy 568 SerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIle 587
Db 2526 TTGAATGCAATTA----- 2537
```

```
Qy 588 GlyLeuGlyTrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeu 607
Db 2538 -----AACAAAT 2543
Qy 608 IleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGly 627
Db 2544 GCTCAACAAGATGATTTAAAGGT-----CGCATCGATCAA----- 2579
Qy 628 ValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn----- 641
Db 2580 -----TCACAGCATTTAAATCAAAATCCAAACAAATTTGTA 2612
Qy 642 -----AspGlnLeuArgAlaThrLeuGlyTyrGln 651
Db 2613 GATGAGGCTAAGGCACCTAATCGTGAATGGATCAATG-----TCA 2654
Qy 652 GlnGluValPheGlyHis-----SerThrAsnGlyPheAspLeuSer 665
Db 2655 CAAGAAATCTCTGCAATGAAGGACGACGAAAGGTAGCACGACTATGTCAATGCAGAT 2714
Qy 666 ThrArgThrLeuGlu-----HisGluIleSerArgSerIleIleGln 679
Db 2715 ACACAACTTAAACAAAGCTGTATGATGAACGGTGTATTAAGGAAACAGCATTGATAAA 2774
Qy 680 AsnGlyGlyTrpAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGln 699
Db 2775 TCGACAGTCAAAACTTAACATCGCAGACAAAGTTATCAAAATTAATGATGCAGTCACCTGCA 2834
Qy 700 AlaProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysPro----- 717
Db 2835 GCTAAGAAAGCA-----TTAAATGCTGAAGAAAGACTT 2867
Qy 718 -----SerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAsp 734
Db 2868 AATAATCGTAAGTCTGAAGCAATTCACAAGA-----TTGGAT 2903
Qy 735 AsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSer 754
Db 2904 CAATTAACACATCTAAACAATGCTCAAAAGACAATTAGCAATCCAAACAAATTAATAATGCT 2963
Qy 755 GlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSer 774
Db 2964 GAAACGCTAAATAAGCATCTCGAGCAATTAATAGACG-----ACT 3005
Qy 775 PheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAla 794
Db 3006 AAATTAGATAATGCAATGGGTGCA-----GTACAACAATATATTGACGACACACCTT 3059
Qy 795 GlyTyrIleThrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGly 814
Db 3060 GGTGTTATCAGCAGCACAAATTCATCATCAATGCAGATGATAATTTGAAA-----GCAAT 3113
Qy 815 GlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGly 834
Db 3114 TATGATAATGCAATTCGAATGCAGCATCATGAG-----TTAGATAAA--- 3155
Qy 835 TyrLeuThrGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMet 854
Db 3156 -----GTGCAAGGTAAATGCAATTCGAAAGCTGAAAGCAGACGACGATTTGAAACAAATAT 3209
Qy 855 LysAspLeuArgLeuAlaValPheGlyAsp-----IleGlyAsnAlaTyrAspLys--- 871
Db 3210 ATCGATGCTCAAAATGCAATTAATGGAGACCAAAACCTTGCATAATGGCAAGATAAAGCA 3269
Qy 872 ---GlyPheThrAsn 875
Db 3270 AATGCGTTTGTAAAT 3284
```

RESULT 21
US-09-815-242-8513
; Sequence 8513, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:

```
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Cart, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8513
; LENGTH: 9477
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; NAME/KEY: CDS
; LOCATION: (1)...(9477)
US-09-815-242-8513
Alignment Scores:
Pred. No.: 0.018 Length: 9477
Score: 131.50 Matches: 184
Percent Similarity: 33.76% Conservative: 135
Best Local Similarity: 19.47% Mismatches: 345
Query Match: 2.78% Indels: 281
DB: 10 Gaps: 48
US-09-914-168-2 (1-919) x US-09-815-242-8513 (1-9477)
Qy 48 AlaIleAsnGlnAlaLys-----AlaGlyAsnProProValLeuLeuThrProGlu 64
Db 4336 GCAGTGAATGACGCTAAACAACTTACATGGTGTATCAAAAATTAGCTCAAGATAAGCAA 4395
Qy 65 GlnIleGlnAlaArgLeu---AsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAla 83
Db 4396 CGTCCACACAGAAACGTTAAATAAAGTTCTTAACTTGAATACA---CCACAACGTTCAAGCA 4452
Qy 84 LeuAsp-----ValValAsnPheAspAspGlnSerProIleSerArgIleGlyGluGln 101
Db 4453 CTTGAAAATCAAAATCAATAATGCAAGCAACTCGTGGCGAAGTAGCACAAAAATTAACGTAG 4512
Qy 102 SerProProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGlu 121
Db 4513 GCACAAGCATTACCAACGAATGAAGCATTCAGTAAAT-----AGCATTTCAA 4560
Qy 122 GluLeuPheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGlu 141
Db 4561 GAT-----CAACAGCAAAACAGACGGGT-----AGCAACTTTTATCAATGAA 4602
Qy 142 -----TyrGlnGlyGluGlnProAsnSerGluValValValPro 154
Db 4603 GATAGGCCACAAAAAGATGTCATACCAACAGCAGCAGTACCAACAGATTTGATTAAAT 4662
Qy 155 ProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsn 174
```


QY 795 GlyTyrIleTrpSerAspAsnProAlaAsnHisValProTyrArgLeuArgPheAlaGly 814
Db 6454 GGTGTTATCAGCAGCAAAATATCATCATGCGAGATGATAATTGAAA-----GCAAAAT 6507
QY 815 GlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGly 834
Db 6508 TATGATAATGCAATGCGAATGCGACACATGAG-----TTAGATAAAA--- 6549
QY 835 TyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPheMet 854
Db 6550 -----GTGCAAGGTAAATGCAATGCGAAGCTCAACGACGAGCAATGCAACAAATATT 6603
QY 855 LysAspLeuArgLeuAlaValPheGlyAsp-----lleGlyAsnAlaTyrAspLys--- 871
Db 6604 ATCGATGCTCAAAATGCAATTAATGAGACCAAAACCTTGCAAAATGCGCAAAAGATAAGCA 6663
QY 872 ---GlyPheThrAsn 875
Db 6664 AATCGGTTTGTGTAAT 6678
RESULT 22
US-08-781-986A-59
; Sequence 59, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31096 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-59
Alignment Scores:
Pred. No.: 0.102 Length: 31096
Score: 131.50 Matches: 181
Percent Similarity: 33.45% Conservative: 118
Best local Similarity: 20.25% Mismatches: 324
Query Match: 2.78% Indels: 271
DB: 7 Gaps: 43
US-09-914-168-2 (1-919) x US-08-781-986A-59 (1-31096)

QY 31 GlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisAspThrAlaIleAsn 50
Db 28877 CAAACTAACAAATCCACCGCTTGATAAAGCACAAAGTTGAACAATTGACACAACTGTTAAAC 28936
QY 51 GlnAlaLys-----AlaGlyAsnProValLeuLeuThrProGluGlnIleGln 67
Db 28937 CAAGCTAAAGATAACCTTACACGCTGATCAAAAACCTTGCACAGCATAAACAACATGCGGTT 28996
QY 68 AlaArgLeuAsnAlaAla---GlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspVal 86
Db 28997 ACTGATTAAATTAATTAATGTTTGAAT---AATCCGCAACCTCAAGCACTT----- 29047
QY 87 ValAsnPheAspAspGlnSerProIleSerArgIle-----GlyGluGlnSerPro 103
Db 29048 -----GAAAGCCAAATAAACAAACGACGACAACTCGTGCGCAAGTAAAC 29092
QY 104 ProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluLeu 123
Db 29093 AAATTAGCT---GAAGCAAAAGCGCTTGATCAGCAATGCAATGCAATGCAATGCAAT 29149
QY 124 PheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGlu----- 141
Db 29150 CAAGATCAACAACAACAGAGATCTGGT-----AGCAAGTTTATCAATGAAGATAAA 29200
QY 142 -----TyrGlnGlyGluGlnProAsnSerGluValValVal----- 153
Db 29201 CCGCAAAAAGATGCTTACCAAGCAGCAGCTTCAAAATGCAAAAGAGATTTTAATTAACCAACA 29260
QY 154 --ProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAla----- 170
Db 29261 GGTANTCCCAACACTCGACAACATCAACAGTACAGCAATTTGACACAAAGCAGTAACTGCA 29320
QY 171 -----ArgLeuPheAsnAsp-----Gly 176
Db 29321 AAAGATAATATACATGGTGCATCAAAAACCTGCTCGTGATCAACAACACAGCAGTACAACT 29380
QY 177 ValAsnLysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThr 196
Db 29381 GTAAATGATGATCCCAACTTAAT-----CATGCACACAACAACAGCATTAACT 29428
QY 197 SerAlaIleGlySerSerHisGlnLysThrGluProTyrTyrAlaAsnIleLysAlaAlaLeu 216
Db 29429 GATGCTATAATGACGCGCTCAAGACAGAGGTTGCAACATGTTCAACACTGCTACT 29488
QY 217 GluAspIleThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGln 236
Db 29489 GAA----- 29491
QY 237 ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArg 256
Db 29492 -----CTTGATCAGCGCATGGAACACATTGAAA 29518
QY 257 AsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGluProValTyrIleAsp 276
Db 29519 AATAAAGTTGATCAAGTGAATACA-----GATAAGCTCAACCAATTAACACT--- 29566
QY 277 TyrArgAlaValGluValArgGlyGlyAlaAspLysAlaPheThrThrValAla 296
Db 29567 -----GAAGCGTCAACTGAT----- 29581
QY 297 AspGluValProLeuLeuIleGlyAspValPheHisGlyLysTyrGluThrLysLys 316
Db 29582 -----AAAAA 29587
QY 317 AsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAspArg 336
Db 29588 GAAGCAGTATGATCAAGC-----TTACAAGCT 29614
QY 337 SerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGly 356
Db 29615 GCAGAAAGCATTTACAGATCCAACTAAT-----GGT 29644
QY 357 ThrGlnTyrArgPheAspGluVal-----ValPhePheThrIleAspProLysThrAsn 374

Alignment Scores:				1			
Pred. No.:	0.0138	Length:	7302				
Score:	131.00	Matches:	181				
Percent Similarity:	32.94%	Conservative:	157				
Best Local Similarity:	17.64%	Mismatches:	346				
Query Match:	2.77%	Indels:	342				
DB:	10	Gaps:	46				
US-S-09-914-168-2 (1-919) x US-09-815-242-4780 (1-7302)							
Qy	29	LeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisAspThrAla	48				
Db	3526	ATTAGTGCACGAATAACCCCTGAATGAATCCTGATACATAATTAACCAAAAGCCAGCCAA	3585				
Qy	49	IleAsnGlnAlaLysAlaGlyAsnProValLeuLeuThrProGluGlnIleGlnAla	68				
Db	3586	GTGAACAGTGGAGTCTGCA-----TTGAACGGTGAATGAA-----	3621				
Qy	69	ArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsn	88				
Db	3622	AAATTAGCAGCAGCAAAACAACTCGAATCAGATATCGTGGTTTGACAGACTTGAAC	3681				
Qy	89	PheAspAspGlnSerProIleSerArgIleGlyGlnGlnSerProProLeuLeuGlyLeuAsp	108				
Db	3682	AATGCACAAACGAACGCGGCAAACTGCTGAAGTGGATCAAGCACCACAAATCTTGACAGCTGC	3741				
Qy	109	MetSerValIleGluGluThrThrProLeuSerLeu-----GluGlu	122				
Db	3742	ACAGCGGCTAAAAATAAAGCAACATCGTTAAACACAGCGATGGGTAATTTGAAACATGCA	3801				
Qy	123	LeuPheAlaGlnGlnSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyr	142				
Db	3802	CTGTCTGAAAGAGTAATACGAACGCTAGTGTCAATTACACAGAT-----	3846				
Qy	143	GlnGlyGlnGlnProAsnSerGluValValValProProThrLeuGluProGluLysPro	162				
Db	3847	---GCGGATCAACCAAAACAACAGCG-----	3870				
Qy	163	GlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArg	182				
Db	3871	-----TATGATCTCGGTTACACAAAGCAGAGCA	3900				
Qy	183	LeuLysAlaLysPheTyrGlnSerGlnSerGlyGluThrSerAlaIleGlySerSer	202				
Db	3901	ATT-----ACTAATGCTAATGCGACCAAC	3924				
Qy	203	HisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGlnGlu	222				
Db	3925	GCGAATGAAACACAA-----GTTCAAGCAGCAGCTAACCCAAATTTGAATCAA---	3969				
Qy	223	SerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAla	242				
Db	3970	GCTAAAATGACTTGAATGGTGAT---ATAAAGTTGCTCAAGCAAAAGAGTCAGCGAAA	4026				
Qy	243	ArgAlaValGlyTyrTyr-----Asp	249				
Db	4027	CGTGCATTAGTTCATATAGTAACTTGAATAATGCGCAATCAACTCGAGCAACTAGTCAA	4086				
Qy	250	IleAsp-----LeuSerIleIleArgAsnSerIleGlyGluVal	262				
Db	4087	ATTGACAAATGCAACGACAGTAGCAGGCGTAACTGCTGCACAAATACTGCTAATGAATTA	4146				
Qy	263	AspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluVal	282				
Db	4147	AATACAGCAATGGGTCAACTT-----	4167				
Qy	283	ArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeu	302				
Db	4168	---CAAAATGGTATAAATGACCAA-----ACACTGTTAAACACAAAGTCAACTTTACA	4218				
Qy	303	IleGlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAla	322				

Db	4219	GATGCTGAC-----CAAGCTAAGAAAGATGCTTTACACAAATGCTGTGTACCAATGCT	4269				
Qy	323	SerAla-----GluHisGlyTyrPheAspGlyArgTrpLeuAspArgSerVal	338				
Db	4270	CAAGGTATTTTAGATAAAGCACACGGTCAA-----AATATGAGAAAGCACAAGTT	4320				
Qy	339	AspValIleLeuProAsp---AsnThrAlaAspValSerLeuIleTyrAspThrGlyThr	357				
Db	4321	GAAGCTGCATTAAATCAAGTAACGACTGCTTAAGAAATGCTTTAAACGGTGCAGCAATGTA	4380				
Qy	358	GlnTyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThr	377				
Db	4381	AGACAAGCA-----AAATCAGATCGGAAGCAAACTTTAGGTACA	4419				
Qy	378	ThrAspProAspLysLeuProValLysArgGluLeuGluGlnLeu-----Leu	394				
Db	4420	TTACACACTTAATTAATGCAAAACAAACAGATTTTAACATCACAATCGAAGGCGCAACA	4479				
Qy	395	ThrValAsn-----MetGlyGluAlaTyr	402				
Db	4480	ACAGTCAACGGTAAATGGTGTAAACGAAAGCACAAGACTTAGATGGTGCATGCCAA	4539				
Qy	403	AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMet	422				
Db	4540	CGATTACAACTCAGCTATAGCAAAATAAAGATCAAACTAAAGCGAGCGAA-----AATTAC	4593				
Qy	423	ValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPhe	442				
Db	4594	ATCGACGCGAGT-----CCAACTAAGAAACAGCATTGTGATAATGCTATCACACA	4644				
Qy	443	GluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluPro	462				
Db	4645	GCTGAATCTTACTTAAATAAGATCAGCGTGGCAATAAAGATAAGCAAGCTCTTCAACAA	4704				
Qy	463	ValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPhe	482				
Db	4705	GCAATTCAAAGTGAACGCTCTACTGAAATGCTTTGAACGGTGACGCGCAACTTACACGC	4764				
Qy	483	SerAlaSerAsnLeuIleGlnAspLysLeuAsnLeu-----	494				
Db	4765	GCTAAACTCAAGCTATACAACTATCGATAACTTCACACATTTTGAATACACCACAAAA	4824				
Qy	495	-----ValAlaAlaLysAlaArgHisLeuTyrAspMetPro	506				
Db	4825	ACAGCATTAAACCAACAAGTGAACGCTGGCAACGCTATCAGGTGTACTGCTAACTGTGAAA	4884				
Qy	507	AspAspArgValLeuAlaIleAsnHisAspGlyValAsnArgSerIleLeu-----	524				
Db	4885	AATAGTGTACATCACTTAAATATGCGATGGATCAATTAAACACAAAGTGTGATCAT	4944				
Qy	525	-----GlyArgIle	527				
Db	4945	GACACAATTGTAGTGTGTAATTACACTAACCGCGAGTCTGTATAGCAAGGTGCTTAT	5004				
Qy	528	SerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluVal	547				
Db	5005	ACTGATGCATATATGCTGCGAAAAACATTGTA-----AATGGTTCGCCCTAATGTGATT	5058				
Qy	548	IleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspVal-----	565				
Db	5059	ACAAATGCACGAGTGTACAGCAGCAACACACGCTGTTAATAATGCTGAAACAGGTTTA	5118				
Qy	566	-----TyrGlnSerLysLysValProLeuTyrVal	575				
Db	5119	AACGGTGCATACAACTTAGCAACTGCGAAAGCAACAACTAAAGATGCTTACGTCAAAATG	5178				
Qy	576	PheValAlaSerAspLysProArgAsp-----GlyGlnIle	587				
Db	5179	ACACATTTATCTGATGTCACAAAAACAAGATTAATTAAGTGTGCTCAAAATGATAGCGCGACACA	5238				
Qy	588	-----GlyLeuGlyTyrGlySerAspThrGlyThrArgLeu-----ValThrLys	602				
Db	5239	GTAAGTGGCTTCAAGTGTCAAGACCAACGCGACAAATCTTGAATGAATGAATCAATCA	5298				

```

QY 603 PheGluHisasnLeuIleasnArgAspGlyTyrGlnAlaGly----- 616
      :::::  :::::  :::::  :::::
Db 5299 CTTGGAATAGTATTGGCAATTAAGATGATGTAAGAGGAGTCAACCATATGTTGATGCA 5358
QY 616 ----- 616
Db 5359 GATGAGATTAACAAATGTCATACATACAGCAGTTACAAATGCTGAATATATCATTAAT 5418
QY 617 ----- 617
      :::::  :::::  :::::  :::::
Db 5419 GCAACGAGTCAGCCGACACTGATTCATGTCAGTACACAGCAAGCATTAACATGAGCAG 5478
QY 623 GluAspLys-----LysGlyValLysLeuTyrAlaThrLys----- 634
      :::::  :::::  :::::  :::::
Db 5479 ACTACACAAATACGGCTTAATGTCGACAAACCTTAGCGAATTAACACAGCAAGCAAGCAGT 5538
QY 635 -----ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
      :::::  :::::  :::::  :::::
Db 5539 GCTAACATCAACCAATTAGTCATTTAATATGCTCA----- 5577
QY 651 GlnGlnGlnValAlaPheGlyHisSerThrAsnGlyPheAspLeuSerThr----- 666
      :::::  :::::  :::::  :::::
Db 5578 AAGCAAGATTTTAATACGCAAGTACAGATGACCAATATATAGCAGTAATCAAGTCAAGT 5637
QY 667 ArgThrLeuGlnHisGluIleSerArgSerIle-----IleGlnAsnGlyGlyTyrAsn 684
      :::::  :::::  :::::  :::::
Db 5638 AAACTAAAGCTGAACATTTGATGATCAAGCAGTGAACGTTTAATACGCAAGTCAAGCAGT 5697
QY 685 ArgThrTyrSerLeuArgTyrArgLeuAspLysLeuTyrThrGlnAlaProProGluThr 704
      :::::  :::::  :::::  :::::
Db 5698 AAA-----GATCAAGTGAACAAGT----- 5718
QY 705 TrpGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGlnAlaLeuLeuAla 724
      :::::  :::::  :::::  :::::
Db 5719 -----GTTAACTTTACAGATGACAGTCCAGAAACCAACAGCATACAC 5763
QY 725 GlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
      :::::  :::::  :::::  :::::
Db 5764 AATGCGGTA-----ACGCGCGTGAATAATATATATATCAAGCAATGCT----- 5808
QY 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
      :::::  :::::  :::::  :::::
Db 5809 -----ACAAATGCGCAACCAATCACCA 5829
QY 765 AlaArgAlaGlyIleSerGlyValTyrSerPhe-----GlyAspAsnAla 779
      :::::  :::::  :::::  :::::
Db 5830 GTTGACAGCAGACCTTTCACTGTAACACTACGAAACAGCGCTTGATGATGAA 5889
QY 780 Tyr-----GlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIle 797
      :::::  :::::  :::::  :::::
Db 5890 GTACAGATGCTAAACCAATGCAACCAACCAATTTATCTAGCTTA----- 5934
QY 798 TrpSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGln 817
      :::::  :::::  :::::  :::::
Db 5935 -----GATACTTAAACCAAT----- 5949
QY 818 SerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuThr 837
      :::::  :::::  :::::  :::::
Db 5950 -----GCACAAACAAAGGCTGCTTACT 5970
QY 838 Gly-----GlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGlyPheMetLys 855
      :::::  :::::  :::::  :::::
Db 5971 GGAACATCATCAACGCGCAACAGTACCTGACAGTACGCAAGCAGCATTAACAAACCGCTGAG 6030
QY 856 AspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsn 875
      :::::  :::::  :::::  :::::
Db 6031 GAATCAATACAGCGAGT-----GTAACCTTGAACAAATAGCTTGAAT-----GATAAA 6078
QY 876 AspThrLysIleGlyAla 881
      :::::  :::::  :::::  :::::
Db 6079 GACACTACACTTGGCAGT 6096

```

```

RESULT 24
US-09-815-242-8898
; Sequence 8898, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8898
; LENGTH: 18846
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(18846)
US-09-815-242-8898

Alignment Scores:
Pred. No.: 0.0552 Length: 18846
Score: 131.00 Matches: 181
Percent Similarity: 32.94% Conservative: 157
Best Local Similarity: 17.64% Mismatches: 346
Query Match: 2.77% Indels: 342
Gaps: 46

US-09-914-168-2 (1-919) x US-09-815-242-8898 (1-18846)
QY 29 LeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisArgThrAla 48
      :::::  :::::  :::::  :::::
Db 5863 ATTAGTCGACAGCAATTAACCTGAATGATCTGATACCAATTAACCAAAACGCCCA 5922
QY 49 IleAsnGlnAlaValAlaGlyAsnProProValIleLeuThrProGlnGlnIleGlnAla 68
      :::::  :::::  :::::  :::::
Db 5923 GTGAACAGTGCAGACTGTCA-----TTGACGGTGATGAA----- 5958
QY 69 ArgLeuAsnAlaValGlyLeuAsnAlaLysProGlnIleAsnArgValValAsn 88
      :::::  :::::  :::::  :::::
Db 5959 AATTTAGCAGCAGCAAAACCAACTCGGAATATGATATCGCTTTGACAGACTTGAGC 6018
QY 89 PheAspAspGlnSerProIleSerArgIleGlyGlnGlnSerProProLeuGlyLeuAsp 108
      :::::  :::::  :::::  :::::
Db 6019 AATGCAACAGCAAGTACGCGCAATATGCTGAAGTGAATCAAGCAACCAATCTTCAGCTGC 6078
QY 109 MetSerValIleGlnGluThrThrProLeuSerLeu-----GluGlu 122
      :::::  :::::  :::::  :::::
Db 6079 ACAGCGGCTAAATAAAGCAACATGTTAAACAGAGCATGGTAATTTGAACATGCA 6138

```

Oy	123	LeuPheAlaGlnGlnSerSerArgThrGlnProAlaGlnValAspGluSerThrLeuGlnPro	142
Db	6139	CTTGCTGAAGAAGGATTAATACGAACGTAAGTCAATTACACAGAT	618
Oy	143	GlnGlnGlyGlnGlnProAsnSerGluValValAlaProProThrLeuGlnProGlnLysPro	162
Db	6184	---GGCGATCAACCAAAACACAACGCG	620
Oy	163	GlyLeuIleLeuArgLeuTyrrAlaArgLeuPheAsnAspGlyValAsnLysValProArg	182
Db	6208	-----TATGATCTCGGGTTTCACAGCAGAGCA	623
Oy	183	LeuLysValAlaLysPheTyrrGlnSerSerGlnSerGlnSerGlnThrSerAlaIleGlySerSer	202
Db	6238	ATT-----ACTAATGCTAATGGCAGCAAC	626
Oy	203	HisGlnLysThrGlnProTyrrAlaAsnIleLysAlaAlaLeuGlnLysPileThrGlnGln	222
Db	6262	GGCAATGAAGAACCA---GTTTCACACACCTAAACCAAAATGGATCAAA	6300
Oy	223	SerAlaMetLysPheLysGlnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAla	242
Db	6307	GGTAAAAATGACTTGATGATGAT---AATTAAGTGTCTCAACCAAAAGCTCAGCGA	636
Oy	243	ArgAlaValGlyTyrTyr-----Asp	249
Db	6364	CGTGATTAAGCTTCATATAGTAACCTGAATATGCGCAATCACTGCACACTAGTCA	6422
Oy	250	IleAsp-----LeuSerIleIleArgAsnSerIleGlyGlnVal	262
Db	6424	ATTGACAAATGCACAGCAGCAGTGCAGCGGCTAAGCTGCACAAATATCTCTAATGAAATTA	648
Oy	263	AspValIleIleHisAspLeuGlyGlnProValTyrIleAspTyrArgAlaValGlnVal	282
Db	6484	AATACAGCAATGGGTCAACTT-----	6504
Oy	283	ArgGlyGlnGlyAlaAspAspLysAlaPheThrThrValAlaAspGlnValProLeuLeu	302
Db	6505	---CAAAATGGTATTAATATGCACAA---AACAGCTTTAAACAACAATGAACTTTTACA	6555
Oy	303	IleGlyAspValPheHisHisGlyLysTyrGlnThrLysLysAsnLeuIleGlnAsnAla	322
Db	6556	GATGGCTGAC-----CAAGGTAACAAGATGCTTACACAATATGCTGTACGAATCT	6606
Oy	323	SerAla-----GlnHisGlyTyrPheAspGlyArgTrrLeuAspArgSerVal	338
Db	6607	CAAGGATATTTAGTAAAGCAACGGGCA-----AATATGACGAAGACACAACTT	6657
Oy	339	AspValIleLeuProAsp---AsnThrAlaAspValSerLeuIleTyrAspThrGlyThr	357
Db	6658	GAAGCTGATTAATCAAGTAACGACGCTCAAGAAATGCTTTAAACGGTCAGCAATGTA	6717
Oy	358	GlnTyrArgPheAspGlnValValPhePheThrIleAspProLysThrAsnGlnLeuThr	377
Db	6718	AGACAAGCA-----AATCGATGATGCAAGCAACACTTAGGTGCA	6756
Oy	378	ThrAspProAspLysLeuProValLysArgGlnLeuLeuGlnLeu-----Leu	394
Db	6757	TTAACACACTTAATATATGCACAAAACAAGATTTAATCATCAATCAAGAACGGCGCAACA	6816
Oy	395	ThrValAsn-----MetGlyGlnAlaTyr	402
Db	6817	ACACTGACCGCTGAATGCTGTTAAACGAAGCAACAAGACTTAGATGCTGCAATGCA	6876
Oy	403	AsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMet	422
Db	6877	CGATTACAAATCAGCTTATACCAAAATAAAGATCAAACTAAAGCGAGGAA-----AATTAC	6930
Oy	423	ValAsnThrGlnLeuValPheProGlnTyrGlnGlnIleGlnAsnAspGlnValSerPhe	442
Db	6931	ATCAGCCAGCAT-----CCAAGTAAAGAAACAGCATTTGATATGCTATACACAA	6981
Oy	443	GlnGlnSerSerSerArgThrGlnProAlaGlnValAspGluSerThrLeuGlnPro	462

Db	6982	GCATGATCTTACTTAATAATAAAGATTCACCGTCCGAATAAGATTAAGCAAGCTTTGAAACA	7041
OY	463	ValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPhe	482
Db	7042	GCAATTCACAAAGTGTAAAGCTCTACTGAAATGCTTTGAACGGTACCGCAACTTACACACC	7103
OY	483	SerAlaSerAsnLeuIleGlnAspLysLeuAsnLeu	494
Db	7102	GCTAAACATGAAGCTTATACAAAGCTATCGATTAAGTACCATTTGAAATACACCAACAAA	7163
OY	495	-----ValAlaIleLysAlaArgHisLeuThrIleuThrAspMetPro	506
Db	7162	ACAGCATTAACAAACAAACAAAGTGAACGGCTCCGCAACGGTGTATACAGGTGTACTGTA	7223
OY	507	AspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeu	524
Db	7222	AATAGTGCTACATCATTAATTAATGCGATGCATCAATTAACAAACAAATTTGCTGATCAT	7283
OY	525	-----GlyArgIle	527
Db	7282	GACACATTTGTAGTACGTAGTGTAATTACACTAACCGGAGCTCTGATAGCAAGGCTCTTAT	7343
OY	528	SerAspAlaValSerAlaValAlaArgAlaIleLeuProAspLysSerGlnAsnGluVal	547
Db	7342	ACGTAGTACATTAATTAAGTGGCAAAACAAATTTGTA-----AATGGTGGCCTATATGATTT	7393
OY	548	IleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaSerPhe	565
Db	7396	ACAAATTCACGACGAGTCTTACACGACCAACACACGCTTTATATATGCTGTAACACAGGTTTA	7455
OY	566	-----TyrGlnSerLysLysValProLeuThrVal	575
Db	7456	AACGGTGATACAAACTTAGCAACTGCCAAGCAACAAAGCTAAAGATGTCATTACGTCAATG	7515
OY	576	PheValIleSerAspLysProArgAsp-----GlyGlnIle	587
Db	7516	ACACATTTATCTGATGCGCAAAAACAAAGTATTACTGGTCAATTAATGTATAGCGCGACACA	7575
OY	588	-----GlyLeuGlyTyrPglySerAspThrGlyThrArgLeu-----ValThrLys	602
Db	7576	GTAAGTCCGCTTCAAGTGTGAAGCAACGACCAACCAATCTGTATATGCAATGATATCA	7635
OY	603	PheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGly-----	616
Db	7636	CTTGCAATATGATATTGGCAATTAAGATGATGTAAAGCGACCAACATATGTTGATGCA	7695
OY	616	-----GlyArgIle	616
Db	7696	GATACAGATAAACAAATTCGATACATATACAGCAGTTACAAATGCTGAAATATCATTAAT	7755
OY	617	-----AlaGluLeuArgLeuSer	622
Db	7756	GCAACGACTCAGCCGACACTTGTATCCATCTGCGATACACAAACAGCATTAATCAAGTGAGC	7815
OY	623	GluAspLys-----LysGlyValLysLeuTyrAlaThrLys-----	634
Db	7816	ACTAACAACAACTGGCGCTTAATGTGTGCACAACCTTAGCAATTAACAAAGCAAGAACGACT	7875
OY	635	-----ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr	650
Db	7876	GCTAAACATCAACCAATTAAGTCACTTTAATTAATAGTCA-----	7914
OY	651	GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThr-----	666
Db	7915	AAGGACATTTAATATACGCAAGTGAACCAAAATTCACCAAAATATTAGCAGTAAATCAAGTG	7974
OY	667	ArgThrLeuGlnHisGluIleSerArgSerIle-----IleGlnAsnGlyGlyTyrPasn	684
Db	7975	AAAACCTAAAGCTGACACATTAAGTCAACGCAATGAGACGTTTAATATCAACGGCAATCCAAAGC	8034
OY	685	ArgThrTyrSerLeuArgTyrArgLeuAspLysLeuThrGlnAlaProArgLeuThr	704

```

Db 8035 AAA-----GATCAAGTCAACAAAGT----- 8055
QY 705 TRPGLNAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAla 724
Db 8056 -----GTTACTTTACAGATGCAGATCCGAAAAAACAAACACCATACAAAC 8100
QY 725 GlyValAlaValAlaHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
Db 8101 AATCGCGTA-----ACTGCTGCTGGAATAATATTATTATCAACCAATATGT----- 8145
QY 745 GlnArgTyrSerLeuGlnValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
Db 8146 -----ACAAATGCGAACCACCAATCCCAA 8166
QY 765 AlaArgAlaGlyIleSerGlyValTyrSerPhe-----GlyAspAsnAla 779
Db 8167 GTTGAAGCAGCACTTCACTGATACACTACGAAACAAACCGTTGATGTTGACAAA 8226
QY 780 Tyr-----GlySerAsnArgAlaHisGlnMetThrGlyIleGlnAlaGlyTyrIle 797
Db 8227 GTAACAGATGCTAAAAACAAATGCAACCAACATTATCTACGTTA----- 8271
QY 798 TrpSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyLysArgIn 817
Db 8272 -----GATTAACCTTAACAAAT----- 8286
QY 818 SerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrIleThr 837
Db 8287 -----GCACAAACAAAGCTGCTGTTACT 8307
QY 838 Gly-----GlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPheMetLys 855
Db 8308 GGAACATCATCAATCAAGCGCACACTGAGCTGAAGTAACCAACCATTCACAAACCGCTCAG 8367
QY 856 AspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsn 875
Db 8368 GAACGGAATACAGCGATG--GTTAACTTGAAAAATAGCTTGAAAT-----GATAAA 8415
QY 876 AspThrLysIleGlyAla 881
Db 8416 GACACTACACTTGCGCAGT 8433

RESULT 25
US-09-815-242-4174
; Sequence 4174, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308

```

```

; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4174
; LENGTH: 3552
; TYPE: DNA
; ORGANISM: Staphylococcus aureus
US-09-815-242-4174

Alignment Scores:
Pred. No.: 0.00539 Length: 3552
Score: 130.50 Matches: 143
Percent Similarity: 35.31% Conservative: 131
Best Local Similarity: 18.43% Mismatches: 287
Query Match: 2,768 Indels: 215
DB: 10 Gaps: 35

US-09-914-168-2 (1-919) x US-09-815-242-4174 (1-3552)
QY 74 GlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnPheAspGlnSer 93
Db 91 GGTCCAAATGGAAGCGGTAAAGTAATATTACAGATGCTTTAAATGG----- 138
QY 94 ProIleSerArgIleGlyGlnGlnSerProProLeuGlyLeuAspMetSerValIleGlu 113
Db 139 -----GTGTTGGCGCAACATCGGCTAA--TCATTACGTGCTCAAAAATGGA 186
QY 114 GluThrThrProLeuSerLeuGluLeuPheAlaGlnGlnSerThrGluMetGlyIle 133
Db 187 GATATTATCTTTCACAGTGCAGACATCCCAAGCTCAAAATTTATTCCTGAAGTACAGTTA 246
QY 134 AsnProAsnAspTyrIleProGluTyrGlnGlyGlnProAsnSerGluValVal 153
Db 247 AGATTAGATATATCTTAAAGCTCAGTGTGATGAA-----AACGAGTTATTGTA 300
QY 154 ProProThrLeu-----GluProGlu----- 160
Db 301 ACAAGAAGATTGTATGGAAGTGTGAAAGAGTACTACATAAATATGACCGTCAACA 360
QY 161 -----LysProGlyLeuIleLysArgLeuTyrAlaArg 171
Db 361 TTAAAGATATTGCCGATTTATTTTATGATTCTGATGGGAAGAACGATATATACATT 420
QY 172 LeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyrGlnSerSer 191
Db 421 ATCTCCAGAGCTAGATTTGATGAATA-----CTAATGCTTAACCAATTGATAGACGT 474
QY 192 GlnSerGlyGluThrSerAla-----IleGlySerSerHisGlnLysThrGluProTyr 209
Db 475 CAATATTATTGAAGAAATCGCTGCTGACTTAATATTAATAAAACGTAACGTAATCATTA 534
QY 210 AlaAsnIleLysAlaAlaLeuGlnAspIleThr-----GlnIleSerAlaMetAspLeu 227
Db 535 AATAAATCTTGACCAACAGAGAAATTAATTAACAGAGTAAACAACTTTATATGATTTG 594
QY 228 AsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyr 247
Db 595 GAAGTCCGCTAGCAACCTGTAAGAGAGAGCGCATATACCTAAA----- 639
QY 248 TyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIle----- 265
Db 640 -----GAATATTAAGACACTTTCGCACCAATGAACATGTCACATTGATGTACA 690
QY 266 IleHisAspLeuGlyLysProVal-----TyrIleAspTyrArgAlaValGlu 281
Db 691 GTCCACGATATTGATCAATATACAAATGCAATAGCAATTAATTAACCTTTAAATGAT 750
QY 282 ValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeu 301
Db 751 TTACAGAGTCAACAAACAAATAAAGAA-----GCTGACAAAGCAACGTTTA 795
QY 301 ----- 301

```

Dh 796 AGCCACAATTCACAATATTAAGCTAAACGTCATCACTTGATTAATGCTGTAATCG 855
Qy 302 -----LeuilegIyAspValPheHISISGLySTyTGLUthrlsYsAsn 317
Dh 856 CTTAATATTCATTAAGTAAAGCTAGCAACCTTTGAAAAATATACGGACAAATTAAT 915
Qy 318 LeuilegIyAsnAlaSerAlaGluHISGLyTrrPheAspGlyArgTrrLeuAspArgSer 337
Dh 916 GTTTAGCAAGACGTAAACAAAATCAATCTGAACCAAAATGACGATATGAGAACAA 975
Qy 338 ValAspValIle-----LeuProAspAsnThrAlaAspValSerLeuileTyrAspThr 355
Dh 976 GAAATTTAATTAAGCCTTTTGAAGAATATATCAATGAGATTCTGAGCTCAACACT 1035
Qy 356 GlyThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspProLysThrAsnGln 375
Dh 1036 -----TATAAGTCTCGAAAGTAAAGTAAACAAAACAA 1065
Qy 376 LeuThrThrAspProAspLysLeuProValLysArgGluLeuGluGlnLeuLeuThr 395
Dh 1066 CTCACACCT-----GTCATTCGTGAATTAAGAACCACTATATGTT 1107
Qy 396 ValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIle 415
Dh 1108 TCACACGACAGCACACGATGAAAAATTGACAGAAATTA----- 1146
Qy 416 AlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIle 435
Dh 1147 ---AACCAATCTATACATTAATGTCA-----GACCAATCGATGTT 1185
Qy 436 GluAsnAsp-----GluValSerPheGluGlnSerSerSerSerArgThrGlu 451
Dh 1186 AACCATATATCTCTTTTAAACATATCTATAGAGAGAGTGAAGGCTAAAAATCAAGA 1245
Qy 452 ProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAsp 471
Dh 1246 CTA-----GATTCTCGATTAGTTGAACTCTTGACCAATTCGAAGATATTCAG 1293
Qy 472 GlyIleLeu-----MetAspIleSerProIle 480
Dh 1294 GGTCAATTAATAAACGACTAAAAAAGATATCAACAGCAACCAAAAGAACTTCTCGTGA 1353
Qy 481 GluPheSerAlaSerAsnLeuileGlnAspLysLeuAsnLeuValAlaAlaValAlaArg 500
Dh 1354 GATTAAGAAATTAATAATATATGAAAAAGACCTCTCATATCAAAAAAGCAAAACGAA 1413
Qy 501 His-----LeuTyr-----AspMetProAspAspArgVal--- 510
Dh 1414 TACGAACAGCAAAATGTATCAAGCATATCGATATACGAAAAAATGAAACCGATATTGAT 1473
Qy 511 ---LeuAlaIleAsnHISAspAsp-----GlyVal----- 519
Dh 1474 ACTTTGGCAACCCAGAGGAAGATATATCTTATTTTTCATATGGCTGCAAAACATATTTTG 1533
Qy 520 -----AsnArgSerIleLeuGly-----ArgIleSerAspAla 530
Dh 1534 AAAGCTAAAAATTAAGATTAAGAGGATATTCATGCTGACGTCGCGAAATTAATTGATGTG 1593
Qy 531 ValSerAlaValAlaArgAlaIle-----LeuProAspGluSerGluGlnVal 547
Dh 1594 CCATCTAAATTAACCTGAGCAATTTGAACAGCATTTAGCTCTTCATTAACAACATGTTATT 1653
Qy 548 IleAsp-----LeuProGluArgThrAlaLeu 556
Dh 1654 GTAGATTTCACAAAAGATGAGCCACCGATTCATTAATTTTAAAGAAACGTAATTTAGGT 1713
Qy 557 AlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeu----- 573
Dh 1714 CGTGCAGACCTTTTACCATTAATGTTATATACAGAGTAGAGTGGAGGACTGATATTAAAG 1773
Qy 574 -----TyrValPheValAlaSerAspLysProArg----- 583
Dh 1774 TCTATTGCTTAAAGAGCAAAACGATTTATATGATTGCTTCGGAAGACGTTAAAGTAGCA 1833

Qy 584 -----AspGlyGlnIleGly----- 588
Dh 1834 CCAGATATTCAAAATATTAATGCGAATTTATAGTAATACGATTAATGCTGATCATTTA 1893
Qy 589 -----LeuGlyTrrPcyLysAspThrGlnTyrArgLeuValThrLysPhe 603
Dh 1894 AAGCATGCAAAATGAAATGTCGACGTCGATTAATATGCAACCTGATATTGTTACT---TTG 1950
Qy 604 GluHISAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGlu 623
Dh 1951 GAAGGTGATATGTAAACCTCGGTGTCTATGACTGGTGGCGCTCGTAAGCTAA 2010
Qy 624 AspLysLysGlyValIleLeuLeuTyrAlaThrLysProLeuSerHISProLeuAsnGln 643
Dh 2011 AGTATTTCTGTCAAAAAGACGAGTGGACA---ACAATGACACACCAATTTGAAGATTAC 2067
Qy 644 LeuArgAlaThrLeuGlyTyrGlnGlnGluValAlaPheGluHISSerThrAsnGlyPheAsp 663
Dh 2068 TTGCCTCAACAGAAATCATTTGCAACCAATTTAAAGAGTTGAAGATAAAGTATCA 2127
Qy 664 LeuSerThrArgThrLeuGluHISGluIleSerArgSerIleIleGlnAsnGlyGlyTrr 683
Dh 2128 TTAAGTGAAGCTGATTTTGAAGAAAGTCAAAAGCATATATACCTTAA----- 2175
Qy 684 AsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGln 699
Dh 2176 GAGCAAGTCATCATTTTGAAGATGAGCTCGATGATTAACATACACA 2223

RESULT 26
US-09-815-242-8027
Sequence 8027, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlson, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OR INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011a
CURRENT APPLICATION NUMBER: US/09/815,242
PRIORITY FILING DATE: 2001-03-21
PRIORITY FILING DATE: 2001-03-21
PRIORITY FILING DATE: 2000-03-21
PRIORITY FILING DATE: 2000-03-21
PRIORITY FILING DATE: 2000-05-23
PRIORITY FILING DATE: 2000-05-23
PRIORITY FILING DATE: 2000-05-26
PRIORITY FILING DATE: 2000-05-26
PRIORITY FILING DATE: 2000-10-23
PRIORITY FILING DATE: 2000-10-23
PRIORITY FILING DATE: 2000-11-27
PRIORITY FILING DATE: 2000-11-27
PRIORITY FILING DATE: 2000-12-22
PRIORITY FILING DATE: 2000-12-22
PRIORITY FILING DATE: 2001-02-16
PRIORITY FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 8027
LENGTH: 3567
TYPE: DNA
ORGANISM: Staphylococcus aureus
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(3567)
US-09-815-242-8027
Alignment Scores:

Pred. No.: 0.00542 Length: 3567
 Score: 130.50 Matches: 143
 Percent Similarity: 35.31% Conservative: 131
 Best Local Similarity: 18.43% Mismatches: 287
 Query Match: 2.76% Indels: 215
 DB: 10 Gaps: 35

US-09-914-168-2 (1-919) x US-09-815-242-8027 (1-3567)

QY 74 GlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValValAsnProAspSer 93
 DB 94 GGTCAAAATGGAAGCGGTAAAGTAATATTACAGATGCTTTAAATG----- 141
 QY 94 ProIleSerArgIleGluGlnSerProProLeuGluLeuAspMetSerValIleGlu 113
 DB 142 -----GTTGGGCGCAACATCGCTAA-----TCATTAGCTGGCTCAAAAATGGA 189
 QY 114 GluThrThrProLeuSerLeuGluLeuPheAlaGlnGlnSerThrGluMetGlyIle 133
 DB 190 GATATTATCTTCAGCTGCAGAACATCGCAAGCTCAAAATTATGCTGAAGTACAGTTA 249
 QY 134 AsnProAsnAspTyrIleProGluTyrGlnGluGlnProAsnSerGluValVal 153
 DB 250 AGATTAGATATCATCTTAAAGCTCAGTGTGATGAA-----AACGAAGTTATTGTA 303
 QY 154 ProProThrLeu-----GluProGlu----- 160
 DB 304 ACAAGAAGATTGTATCGAAGTGTGMAAGTGAAGTACTACATTAATATGACCGTCAGAA 363
 QY 161 -----LysProGlyLeuIleLysArgLeuTyrAlaArg 171
 DB 364 TTAAGATATTCCGATTTATTATTAGATTCTGATTTGGGAAGAGCATATAGCATT 423
 QY 172 LeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyrGlnSerSer 191
 DB 424 ATCTCCCAAGTGTAGTGTGTAATA-----CTAATGCTAAACCAATTGATAGACT 477
 QY 192 GlnSerGlyLysThrSerAla-----IleGlySerSerHisGlnLysThrGluProTyr 209
 DB 478 CAATATTTCAGAAATCGCTGCTGACTTAATATTAAGTAAGCTAAAGCTAAATCATTA 537
 QY 210 AlaAsnIleLysAlaLeuGlnAspIleThr-----GlnGlnSerAlaMetAspLeu 227
 DB 538 AATTAACCTTGACCAACAGAGATTAATTAACGAGCTAGACATTTATATGATTTG 597
 QY 228 AsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyr 247
 DB 598 GAAGTTCGCTAGAACCTCTAAAGAGAGGAGCGACTATAGCTAAA----- 642
 QY 248 TyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIle----- 265
 DB 643 -----GAATATAAGACACTTTCGCACCAATGAACATAGTAGCATTTGTTAGTACA 693
 QY 266 IleHisAspLeuGlyLysProVal-----TyrIleAspTyrArgAlaValGlu 281
 DB 694 GTGCACGATATTATCATTAATTAACAAATGACAAATAGATAGTCAACCTTTAAATGAT 753
 QY 282 ValArgGlyLysGluAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeu 301
 DB 754 TTACAAGGTCAACAGCAATTAAGAA-----GCTGACAGCAACAGTTTA 798
 QY 301 ----- 301
 DB 799 AGCCAAATATTCACAAATTAAGTAAGTAAGTATCATCACTGATATGATGTTGATCG 858
 QY 302 -----LeuIleGlyAspValPheHisIleGlyLysTyrGluThrLysLysAsn 317
 DB 859 CTTAATATTATCAATTAAGTAAGTACGAGAGCCCTTGAATAATATAGCGGCAATTAAT 918
 QY 318 LeuIleGluAsnAlaSerIleGluHisGlyTyrPheAspGlyArgThrLeuAspArgSer 337
 DB 919 GTTTTGAAGAACGTAAAGAAATCAATCTGAACAACATGCGCATATVGAAGAACAA 978

QY 338 ValAspValIle-----LeuProAspAsnThrAlaAspValSerLeuIleTyrAspThr 355
 DB 979 GAAATTTTAATGAGCTTTAGAAATATATCAATGAGATTTCTGAAGCTCAAGTACT 1038
 QY 356 GlyThrGlnTyrArgPheAspGluValAlaPheThrThrIleAspProLysThrAsnGln 375
 DB 1039 -----TATTAAGTCTCGAAAGTAACAAAGAA 1068
 QY 376 LeuThrThrAspProAspLysLeuProValLysArgGluLeuGlnGlnLeuThr 395
 DB 1069 CTCACAGCT-----GTCATTGCTGAATTAAGAAACAACACTATATGCTT 1110
 QY 396 ValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIle 415
 DB 1111 TCAGACGAGACACGACGATGAAATTTGAGCAATTTAA----- 1149
 QY 416 AlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluIle 435
 DB 1150 ---AACCAATATCTATACATTAATGCTCA-----GAGCATTCAGATGCTT 1188
 QY 436 GlnAsnAsp-----GlnValSerPheGluGlnGlnSerSerSerArgThrGlu 451
 DB 1189 AACATATATTCGTTTAAACCATCTATTAAGAAAGATGAGGCTAAATAATCAAGA 1248
 QY 452 ProAlaGlnValAspGlnSerThrLeuGluProValIleGluThrValGluLeuThrAsp 471
 DB 1249 CTA-----GATTCGATTTAGTTGAAGCTTTTGACCAATTCAAGATTTACG 1296
 QY 472 GlyIleLeu-----MetAspIleSerProIle 480
 DB 1297 GGTCAAAATAAAGACTAAAGAAATATCAACAGACCAACAAACCTTTCGCTGTA 1356
 QY 481 GluPheSerAlaSerAsnLeuIleGlnLysLysLeuValAlaAlaLysAlaArg 500
 DB 1357 GATTAAGAAATTAATAATATATGAAAGAGCTCAGTATACAAAAAGCAAAAGCA 1416
 QY 501 His-----LeuTyr-----AspMetProAspAspArgVal-- 510
 DB 1417 TACGAGAGAAATTTGTATCAAGCATATTCGATATACCGAAATGAAGAAACGATTTGAT 1476
 QY 511 ---LeuAlaIleAsnHisAspAsp-----GlyVal----- 519
 DB 1477 AGTTTGCACAGCAAGAGGAAGATATATCTTATTTTCAATGGCGTCAACATATTTTG 1536
 QY 520 -----AsnArgSerIleLeuGly-----ArgIleSerAspAla 530
 DB 1537 AAACCTAAATTAATTAAGATTTAAAGGATTTATCATGTGTCAGTTCGGAATTTATGATGTG 1596
 QY 531 ValSerAlaValAlaArgAlaIle-----LeuProAspGlnSerGluAsnGluVal 547
 DB 1597 CCATCTAAATTAATCACTGAGCAATTTGAACACAGCATTAGTGCTTCATTAACACATGTTAT 1656
 QY 548 IleAsp-----LeuProGluArgThrAlaLeu 556
 DB 1657 GTAGATTTCAGAAAAAGATGACCGCCAGCTATTCATTTTAAAGAAACGTAATTTAGCT 1716
 QY 557 AlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeu----- 573
 DB 1717 CGTCGACGTTTTTACCATTAATGTTTATACAGAGTAGAGTGTAAGCAGCATATTAAAG 1776
 QY 574 -----TyrValPheValAlaSerAspLysProArg----- 583
 DB 1777 TCTATTGCTAAAGAGCAACGATTTATATGTTATGCTTCGGAAGCAAGTAAAGTACGA 1836
 QY 584 -----AspGlyGlnIleGly----- 588
 DB 1837 CCAGAAATATCAAAATATATATTCGGAATTTATAGTAATATACGATTTATTCGATCATTTA 1896
 QY 589 -----LeuGlyTyrPheAspThrGlyTyrArgLeuValThrLysPhe 603
 DB 1897 AACGATGCAATGATTTGGCAGTCGCTTAATATATGCAACCTCGATTTGTTACT---TTG 1953
 QY 604 GluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLysSerGlu 623

Oy	43	ProAlaHisAspThrAlaIleasnGlnAla	-----	5
	111	111	111111	
Db	331	CCTGACAAACAAAAATGCGTATTAACAGGCGACAGCAAAAGCTGAAGCATTTATTGAATAA	3	
Oy	53	LysAlaGlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaArgLeuAsnAla	7	
	111111	111	111	1111111111
Db	391	CAAAAGGTACTAATGAATCAACACAGCAAGTTGAAGACTACTAATGAAGTGAACCA	4	
Oy	73	AlaGly	-----LeuAsnAlaLysProGlnSer	8
	111	1111111111	111111	
Db	451	GGCAAAACAGCATTTAAATGGTATATGACAAATTTGGCAAAATGGCAAAACAAACAGCAAAACAA	5	
Oy	82	GlnAlaLeuAspValValAsnPheAspAsp	-----	9
	111	111111	1111111111	
Db	511	CAATTGGCGCAACTTAAACACACTTAATATGATGCACAAAACAAATCATTTGAAGCAATTT	5	
Oy	92	---GlnSerPro-----IleSerArgIleGlyGlnGlnSerProProLeuGly	1	
	1111111111	1111111111	1111111111	
Db	571	ACACAAGGCGCACCTTGTTACAGAGTGCCTACGATTATATCAAAAAGCAACAAAGCTTGAT	6	
Oy	107	-----LeuAspMetSerValIleGlnGlnThrThrProLeuSerLeuGln	1	
	111111	111111	1111111111	
Db	631	CATGGCATGGAATTAATTAAGAATAGCTGGCCGATATCAAAAGACATTAAGCGTGTGA	6	
Oy	122	GluLeuPheAlaGlnGlnSerThglUmetGlyIleAsnProAsnAspArgIleProGln	1	
	111111	111111	1111111111	
Db	691	GATTAT-----CATGATGCACACTGGCCAA-----AGACAAAATGCTATTAACCAAGCT	7	
Oy	142	TyrGlnGlyGlnGlnProAsnSerGlnValValProProThrLeuGlnProGluLys	16	
	1111111111	1111111111	1111111111	
Db	739	GTAACAGCTGCTAATATATCATTAATCAACATACATGCGCTCGATGATCA	7	
Oy	162	ProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyAlaLysValPro	18	
	1111111111	111	111111	
Db	793	-----GATCATTTAAT	80	
Oy	182	ArgLeuLysAlaLysPheArgInsSerSerGlnSerGlyThrSerAlaIleGlySer	20	
	111	111111	111111	
Db	805	-----GCTGCAACGACACAAAGTGAATAT	82	
Oy	202	SerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGlu	21	
	111111	111111	111111	
Db	829	ACGAACCTTGCAATTAATGATGCTGATGAAAACTTCGACAGCTAACAACAACAAACAAAC	88	
Oy	218	-----AspIle	21	
Db	889	AGACTGCATCAATTACATCATTTGAATAATGCCCAAAACCAACAGTTCAATCAACAATTT	94	
Oy	220	ThrGlnGlnSerAlaMet---AspLeuAsnGlySerIleProArgLeuArgIleThrAla	23	
	111111	111	1111111111	
Db	949	ACGCATCAATCTGATATTTCTCCAGCTTAATAGTGCAC-----AACCAAAACACA	99	
Oy	239	LeuValAlaAlaArgGlnAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer	25	
	1111111111	111111	111	
Db	997	GAATCTTTAAATACATGCGATGGGT-----AACTTA	103	
Oy	259	IleGlyGlnValAspValIleIleHisAspLeuGlnGluProValTyrIleAspTyrArg	27	
	111111	111111	1111111111	
Db	1027	ATTATATGCGATT-----GCAGATCATCA	109	
Oy	279	AlaValAlaGlnValArgGlyGlu-----GlyAlaAsp---AspLys---AlaPheThr	29	
	1111111111	111111	111111	
Db	1051	GCGCGTTGAACAAACGTTGATCACTCATATGCTGATCTGATTAACAACAACTGCTTAAT	111	
Oy	294	ThrValAlaAspGlnValProLeuLeuIleGlyAspValPheHisGlyLysTyrGlu	31	
	1111111111	111111	111111	
Db	1111	ACAGCGGTAAATGAAGACGACGACGATGTTAAACAACAACAACTGCTCAAAATGCGAACCA	117	
Oy	314	ThrLysLysAsnLeuIleGlnAlaAlaSerAlaGlnHisGlyTyrPheAspGlyArgTyr	33	
	1111111111	111	111	
Db	1171	ACAGAA-----GTGAACAACAGCTATTCTCAAA-----	119	


```

Db 1198 GTTCAAAACAACACTTCAAGCGTTAAATGAGACCATTAATTACAAAGTTGCT-----AAA 1251
Oy 354 AspThrGlyThrGlnThrArgPheAspGluValAlaPhePheThrIleAspProLysThr 373
Db 1252 ACAAAATGCGACGCA-----GCAATTGCTTTTAA 1284
Oy 374 AsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGlnLeu 393
Db 1285 AGC-----TTAAATGATCCCTCAAAAACAGCATTTAAAGACCAAGTTACAGCTCAACT 1338
Oy 394 Leu-----ThrValAsnMetGlyGluAla 401
Db 1339 TTGATTAAGTCACTTCAATTAATGAACAAATGCGAATACGCTTAACCAAGCA----- 1392
Oy 402 TyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsn 421
Db 1393 -----ATGATGCTTTAAAGCAGACGATTCACAGATTAACGCAACACTAA----- 1437
Oy 422 MetValAsnThrGluIleValPheProGluArgGlnGlnIleGlnAsn---AspGlnVal 440
Db 1438 ---GCAATATGCAAAATATATCAACAGACGCAACCAACCAAACTATGATCAACT 1494
Oy 441 SerPheGlnGlnSerSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu 460
Db 1495 GTTCAACCCCGCAATATATATATATCAATGAACAACTGCAACATTAATATATATGCGATT 1554
Oy 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMet---AspIleSerPro 479
Db 1555 AATCAACAGCAGCAACTGTGAATACACAGCAACAGCAGCATTAACGTGCGATGTC----- 1608
Oy 480 IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAla 499
Db 1609 -----AAGTTACAA----- 1617
Oy 500 ArgHisLeuThrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyVal 519
Db 1618 -----AATGATTAAGATCATGCT 1635
Oy 520 AsnArgSerIle-----LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAla 537
Db 1636 AAGCAAAACGGTTAGTCACTTAATAGCATCTTAACATGCACAAAACATATATGAAGATACG 1695
Oy 538 IleLeuProAspGlnSerGlnAsnGluValIle-----AspLeuProGluArgThrAla 555
Db 1696 TTAATGATAGTGAACAACACTAGAACAGAGTTAGCAAGATTTGACTGAGCAACAGCA 1755
Oy 556 Leu-----AlaAsnArgLysThrProAlaAspValTyrGlnSerLysLys 570
Db 1756 TTGATCACTTATGATGATGATTCACAAAGATATGCTGCAAAAGATGCAACACGTCGC 1815
Oy 571 ValProLeuThrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGly 590
Db 1816 AGCAGTGCATATGTCATGCAAGCAACCGAATTAACCAATCTATGATGAACAGCT--- 1872
Oy 591 TrpGlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArg 610
Db 1873 -----CAAAATGCTGAGTATCATGTCAGAGATTAAATATCCAACTATCATATAA 1923
Oy 611 AspGlyTyrGlnAlaGlyAlaGluLeu----- 619
Db 1924 GGTAAATGATATCACTGACACTGACAGCAAGTATATCATCTAAAAATGCTATGATGCTGTT 1983
Oy 620 ---ArgLeuSerGlnAspLysGlyValLysLeuThrAlaThrLysProLeuSerHis 638
Db 1984 GAACGATTACTCAAGTAAAGCAA-----ACTGCTGCAAAATTCCTTAATATAT 2031
Oy 639 ProLeuAsnAspGlnLeuThrArgAlaThrLeuGlyTyrGlnGlnGlnValPheGlyHisSer 658
Db 2032 -----TTAGATCAATTAACACGAGCT-----CAACAAACAAAGCGCTGAAGAAATCAA 2076
Oy 659 ThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHisGlnIleSerArgSerIleIle 678

```

```

Db 2077 ATTAAT-----AATGCAACAACCTCGTGGCAAGTACACAAAAATTAAC 2121
Oy 679 GlnAsnGlyGlyTrpAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuThr 698
Db 2122 GAA-----GCACACAGCTTTAAACCAAGCAATGGAAGCTTTAGGTAT 2163
Oy 699 GlnAlaProProGluThrTrpGlnAspLeuProValAspPheValAsn---GlyLysPro 717
Db 2164 AGCATTAAGATCAACACAGCAACGCGGTAGCAGTTATATCAATGAAGATTAACCA 2223
Oy 718 SerGlnGlnAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuVal 737
Db 2224 CAAAAGATGCTTACCAAGCAGCAGCTCAAAATGCAAAAGATTAATTAATCAAACTAAC 2283
Oy 728 AsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuVal 757
Db 2284 AATCCAAAGCTTGATTAAGCACAA-----GTT 2310
Oy 758 SerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp 777
Db 2311 GAACATTTGACACACAGCTGTTAAACCAAGCTTAAGATTAAC-----CTACACGCTGAT 2361
Oy 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThr 789
Db 2362 CAAAACCTTGACAGCATTAACACACATGCGGTACT 2397

RESULT 28
US-09-954-456-1807
Sequence 1807, Application US/09954456
Patent No. US20020115057A1
GENERAL INFORMATION:
APPLICANT: Young, Paul
TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
FILE REFERENCE: 689290-76
CURRENT FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: US/60/233,617
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US/60/234,052
PRIOR FILING DATE: 2000-09-20
PRIOR APPLICATION NUMBER: US/60/234,923
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,134
PRIOR FILING DATE: 2000-09-25
PRIOR APPLICATION NUMBER: US/60/235,637
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,638
PRIOR FILING DATE: 2000-09-26
PRIOR APPLICATION NUMBER: US/60/235,711
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,720
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,840
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: US/60/235,863
PRIOR FILING DATE: 2000-09-27
NUMBER OF SEQ ID NOS: 2276
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1807
LENGTH: 3336
TYPE: DNA
ORGANISM: Homo sapiens
US-09-954-456-1807

Alignment Scores:
Pred. No.: 0.00781
Score: 128.50
Percent Similarity: 33.76%
Best Local Similarity: 19.89%
Query Match: 2.72%
Indels: 239
Gaps: 46

```

US-09-914-168-2 (1-919) x US-09-954-456-1807 (1-3336)

Oy	95	IleSerArgIleGlyLysInserProPheLeuSylLeuAspMetSerValIleGluGln	114
Db	342	ArCTGTGCAGTGGGAATTCATCATCCGCCGCTTTTGCAATC---TTTGCTGTGGACAAAAC	398
Oy	115	ThrThrProLeuSerLeuGluGlnLeuPheAlaGlnGlnSerThrThiGluMetCylLeuAsn	134
Db	399	ACTGGAGATATTAAACAATAACACTATAGTCGACC GGAGAAACT-----	443
Oy	135	ProAsnSpryTrlLePro-----GluTYrGlnGlyLysInProAsnSerGlu	150
Db	444	CCAAGCTTCTGTATCACATGTCGGGCTTAATGCCCAAGA-----	485
Oy	151	ValValAlaProPheThrLeuGlnProGlyLysProGlyLeuIleLYsAlaGlyLeuTYra	170
Db	486	-----CTAAGTAGAGAAACCA---CTTATATACACGGTTAAAAATT	524
Oy	171	ArgLeuPheAsnAspCly-----ValAsnLYsValProArgLeuLYsAlaLYsPheTYr	188
Db	525	TTCGATATTAAATATATATCTCCACGATTTTTCACAAACAATTTTCATCGGGTGAATTTGAA	584
Oy	189	GlnSerSerInserCylGlnThrSerAlaIleGlySerSerHisGlnLYsThrGluPro	208
Db	585	GAATAATAGTCCCTCAACACTCACTGGTATGATACTAAATGCGACAGATGCAGATGAACAA	644
Oy	209	---TYrAlaAsnIleLYsAlaLeuGlnLYsAspIleThrGlnGlnSerAlaMetLYsLeu	227
Db	645	AACCACTTAATTCATAAATTCTCTCAAAATTCGTCTCAGAACACAGA-----	695
Oy	228	AsnGlySerIleProArgLeuArgIleThrAlaLeuValAlaAlaArgAlaValGlyTYr	247
Db	695	-----	695
Oy	248	TYrAspIleAspLeuSerIleIleArgAsnSerIleGlyLysValAspValIleIleHis	267
Db	696	---GGCACACCCACTGTCTCTCTTAAGCAGAAACACTGGGAATCCGTAATTGACCAAT	752
Oy	268	AspLeuGlyLysProValTYrIleAspTYrArgAlaValCylValArgGlyLysCylala	287
Db	753	TCTCTTGAACCGAGAACAACTAGCAAGCAAGTATGCTGTGTT---CTGAGTGTGCAGACAA	809
Oy	288	AspAspLYsAlaPheThrThr-----ValAlaAsp	297
Db	810	GATGCGAAGCAAGCATCAACTCATGTGATGATATTAAGTGAACAGATGTCAACGAT	869
Oy	298	GluValProLeuLeuIleGlyAspValPheHisIleGlyLYsTYrGlnThrLYs---	315
Db	870	AAC TTC CCAAT G-----TTT TTA GAG CTC CAT GAT CA C AG CAT AT TGA A	914
Oy	316	LysAsnLeuIleGluAsnAlaSerAlaGlnIleGlyTYrPheAspGlyArgTrlPheAsp	335
Db	915	GAATAATATTTTAACTTCTGAATTACTT-----CGATTTCAACTA	953
Oy	336	ArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTYrAspThr	355
Db	954	ACAGATTTCGATGAAGACTACACAGATTAATGGCTTCGACGT---TATTCTTAACTCT	1010
Oy	356	GlyThrGlnTYrArgPheAspArgValAlaPhePheThrIleAspProLYsThAspGln	375
Db	1011	GGCAATGAAGAATTGCTTTGAAATA-----CAAACGTATCCAGAACATTAAGAA	1061
Oy	376	LeuThrThrAspProAspLYsLeuProValLYsArgGlyLeu---LeuGlnGlnLeu	394
Db	1062	-----GGCATCTCGAAGACGTGGAGGCTGTACATGATATGAAACAACATTCAA	1106
Oy	395	ThrValAsnMetGlyLysLYrArgAsnLeuGlnAlaValArgAlaLeuSerAspLeu	414
Db	1107	AGCGTGAACCTTATGCTTCTCTGCAAA-----AACAAAGCTGAATTTCCACAAATCA	1157
Oy	415	IleAlaThrArgTYrPheAsnMetValAsnThrGlnIleValPheProGluArgGlyGln	434
Db	1158	GTTATCTCGATAC---GGAATTTCAGTCAAACCCACAGTCACAATT-----CAG	1202

QY 435 ILeGlnsN-----AspGlnValSerPheGlnGlnSerSerSerArgThrGluPro 452
 Db 1203 GTAATAATGAAAGAGAAAGAAATGGATTCCTGGCTCGCTCCAGACA----- 125
 QY 453 AlaGlnValAspGlnSerThrLeuGlnGluProValILleGluThrValGluLeuThrAspGly 472
 Db 1251 -----TTTACTGTGGCAAAAGGCATTAAGTACCAAAAATTGGTGGATTAT 129
 QY 473 ILleuMetAspIleSerProILleGluPheSerAlaSerLeuILleGlnAspIleu 492
 Db 1286 ATTCCTG-----GGAACTATTCAGCCATCGATGAGCAACT 133
 QY 493 AsnLeuValAlaIleAlaValSerILleArgHisLeu-----TyrAspMet 505
 Db 1332 AACCAACCTCCCTCAATATGTCATATATGTCATGGACGTAAACGATGGTGATACCTAATAG 1391
 QY 506 ProAspAspArgValLeuAlaILleAsnHisAspArgIleValAsnArgSerILleLeuGly 525
 Db 1392 ATTAGTTCAAAACCTCGTAATCAAAATTGTGCAAAATATGAACCGAGATTCTACTTTC 1455
 QY 526 ArgIleSerAspAlaValSerILleValAlaArgValIle----- 538
 Db 1452 ATAGTTATACAAAACAATCACACAGCTGAGGTTCTGGCCATATGATGAATACAGGGTAAACT 1511
 QY 539 -----LeuProAspIleSerGlnGlnGlnValILleAspLeu 550
 Db 1512 TCTACAGCCACGGTATATGTTAGAGTACCAGATTTCATACAAATTCCTCCACACAGCTGTC 15711
 QY 551 ProGlnArgThrAlaLeuAlaAsnArgGlyThrProAlaAspValTyrGlnSerIleuLys 570
 Db 1572 CTCGAAAAAGATGCAGTTTGCAAGT---TCTTCACCTTCCTGGTCTTCTCCCGCTGAGACA 16288
 QY 571 Val-----ProLeuTyrValPheValAlaSerAspIleProArgAsp 584
 Db 1629 CTGAAATATAGATACACTGGGCCC---TATACATTTGCACTGGAAAGATCAACT--- 1679
 QY 585 GlyGlnILleGlyLeuGlyTyrGlySerAspThr-----GlyThrArgLeuValThrLys 602
 Db 1680 GTAAAGTTCCCTGCCGTATGGAGATACACAACCCATAGTCACTCCGGCCCTCTCAGA 1739
 QY 603 -----PheGlnHisAsnLeuILeAsnArgAspGlyTyr 613
 Db 1740 GCCCAGAAACAGATACCTCTCGAGATATACCAATCTCCCTGGTACTTACAGACAGTAC 1799
 QY 614 GlnAlaGlyAlaGluLeuArgLeuSerGln-----AspLysLys 626
 Db 1800 AACCAATCGGTGTGAGATGCCACGACGCTTGACACTGAAAGTCTCTCAGTGTGACACAGG 1859
 QY 627 GlyValILleLysLeuTyrAlaThrLysProLeuSerHisPro----- 639
 Db 1860 GGCATC---TGTGCAACTTCTTACCCACACACAGCCCTGGAGCCAGTATGCAAGCGC 1916
 QY 639 ----- 639
 Db 1917 CACTCAGGAGCGCTGGGCGCTGCCGCCATCGGCGCTGTCTCTGGTCTCTGCTCTGCTG 1976
 QY 640 ---LeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGlnValPheGlnHisSer 658
 Db 1977 CTGTGGGCCCCCTCTCTGCGTGGACCTGTGACTGTGGGCGAGGTTCTACTGGGGGAGTG 2036
 QY 659 ThrAsnGlyPhe-----AspLeuSerThrArgThrLeuGlnHisGlnHisLeuSer 674
 Db 2037 ACAGGTGGTTTATTCACCACTGCTGATGGCTCAGAAAGAACAAATT---CATCAG----- 2087
 QY 675 ArgSerILleILeGlnsGlnGlyTyrAsnArgThrTyrSerLeuArgTyrAlaGluAsp 694
 Db 2088 -----TGGGGA-----ATTGAAGAGGCCATCTCTGAAGAC 2117
 QY 695 LysLeuLysThrGln-----AlaProProGlnThrThrPrlGlnsLeuProValAspPhe 712
 Db 2118 AAGAAATACCAAAATTTGTGTGGCTCTCTCTCAACA----- 2153

Oy	331	GLYVRGTRPLEUASPRGSRVERVALSPVALILEUENPROASP-----	344
Db	2958	GGCGCATATCATCTTTAAGTCC-----ATTTCCTGCTATTATCCCTACGAGATTGG	3001
Oy	345	-----AsnThrAlaAspValSerLeuIleTyrAsp-----	354
Db	3009	TGGACACAGACCATTGAAGAAAAAAGATATTTGGCGTGTGGATTATATGCAATAAGTT	3066
Oy	354	-----	354
Db	3069	GGCGGATTCGGATTACTGCACCTTTTGGATATTTAAAAAGCTTGCTTAAGTAT	3128
Oy	355	-----ThrGLYThrGLNtyrArgPheaspValValPhePheThrIleAspPro	371
Db	3129	GGGACTAAGCGTGGATTTCCTATCTCATATGAGAGATATT-----ATCACGCCA	3176
Oy	372	LysThrAsnGLNLeuThrThrAspProAspLysLeuProValLysArg-----Glueu	389
Db	3177	AAAGACAAACAAAAAATGTTGGAAAAAGCCAAAGTACAGGTTAAAAAATCCAGCAACAA	3233
Oy	390	LeuGLNGLNLeuLeuThrValAsnMetGLYGLNValTyrAsnLeuLlnAlaValArgAla	409
Db	3237	TACGATCAAGGGCTGCTCCTACGACCA--GAGCGTTAC-----	3277
Oy	410	LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGLNValPhe	429
Db	3273	-----AATTAAGATCATTTGACACT-----TGACATCAAGTCAATGACAAATGAGTAA	3320
Oy	430	ProGLNtyrGLNGLNLeuAsnAspGLNValSerPheGLNGLN-----	444
Db	3321	GAAATGATGACCCGCTATTCGGCAAGATAAAGAGCGCTTAACCTATATTTATGATGCA	3380
Oy	445	SerSerSerSerArgThrGLNProAlaGLNVal-----	455
Db	3381	GATTACCGCGCGAAGGGGTAGCGCGCAATCCGTCACGCTTACAGCATGAGGGCTCT	3440
Oy	456	-----AspGLSerThrLeuGLN-----ProValIleGLNThrValGLNLeuThr	470
Db	3441	ATGACAAAGCCGGACGGCAGATCATTTGAACCCCATTAATTTCTAACCTTAAAGAGGG	3500
Oy	471	AspGLYIleLeuMetAspIleSerProIleGLNPhSerValSerAsnMetIleGLNAsp	490
Db	3501	TTGAATGCTCTTAAGATCACTTCATTCACAGCATGCGCCATGAAGAGCGCTTACGGATACA	3560
Oy	491	LysLeuAsnLeuValAlaAla-----LysAlaArgHisLeuTyrAspMetProAspAsp	508
Db	3561	GGCGTAAAAACAGCCCAATGCGGGGTATTTGACCAGCAACCATCATTTGATGATT--TCGCAA	3617
Oy	509	ArgValLeuAlaIleAsnHisAspAspGLYValAsnArgSerIleLeuGLNArgIleSer	528
Db	3618	AATGTCAAGCTGGTGTCTGTGATGTTCCGGCACCGCATGAAGAGGATT-----GAATCAGC	3671
Oy	529	AspAlaValSerAlaValAlaArgAlaIleLeuProAspGLNserGLN-----	544
Db	3672	GAT---ATTCCGGTGGGAGTACGATGATTGAACCTTTGAAGAAGACGATTTTGGGGCGC	3728
Oy	545	-----AsnGLNValIleAspLeuProGLNtyrThrAlaLeuAlaAsnArgLysThr	561
Db	3729	GTTTATTATACAAATATGATGATTCATCCATTTACGAATGAATCTTCCTTATGCGGACACT	3788
Oy	562	ProAlaAspValTyrGLNserLysLysVal-----ProLeu	573
Db	3789	TTGATTGATGAAAGGGTCTCTAAAAAGTGGTGTGAAGCCGGGATTTAATCATTCATGATC	3848
Oy	574	TyrValPheValAlaSerAspLysProArgAsp-----GlyGLNLeuGLY	588
Db	3849	CGCACCCCACTTAATTCTGTAAGCGCCAAAGGGCGTGTCCGGAATGCTATGGCTTGAAAT	3908
Oy	589	LeuGLYTrpGLY-----	592
Db	3909	TTGGGGAAGCGCAAGTGTAGTTATCCGGGTGAAGCGGTGGCGCTGTGTAAGCGCGACATCT	3968

QY 593 ---SerAspThrGlyThrArgLeuValThrIlysPheGluHis----- 605
 Db 3969 ATTGGGAGACCGCTGGAAAGCCAGCAGCTCACTTTAAAGACCTTCCATGTGGCGGGAGACAGCGAC 4022
 QY 606 -----AsnLeuIleAsnArgAspLeuTyrglnAlaIleValIleGlu 618
 Db 4029 AGGAGCTCAGAGTACAGCGCGGAATCGTATGCGACGCAAGAAAGAAAGTGTGTGGCTTTTAAAC 4088
 QY 619 LeuArgLeuSerGluAspIleValGlyGluValIleLeuThrLeuThrIleHis----- 634
 Db 4089 CTTAGAGCTTACAGGAATAAAGAGGGTAAACATTATGCGTATACCGCGCTTAACGCTTCT 4148
 QY 635 -----ProLeuSerHisProLeuAsnAspIleValArgAlaThrLeu 648
 Db 4149 ATTTAGTGTAGAGCCTAAAGCTAAATGATAAAGCCGCTTTGATGGAGGAATTACGATGAAAG 4208
 QY 649 GlyTyrglnGlnGluValPheGluHisSerThrAsnGlyPheAsnLeuSerThrArgThr 668
 Db 4209 GTTTATGAAGAAGTCGTTGACG----- 4232
 QY 669 LeuGlnHisGluIleSerArgSerIleIleGlnAsnGlyGlyTrpAsnArgThrTygSer 688
 Db 4233 -----GTCAAAATAAGCGAGATCAAGAGCTAAATTTGTT 4265
 QY 689 LeuArgTyArgLeuAspIleValGlySerThrGlnAlaProProGluThrTrpIleAsnLeu 708
 Db 4266 TTAAGC---ACAAGCGATATTGTCAAG-----CCAAGCGAA----- 4298
 QY 709 ProValAspPheValAsnGlyIlysProSerGlnGluAlaLeuLeu-----AlaGly 725
 Db 4299 ---TTAGCCGGCGCTGGCGGTAGAGATTGAAGGGGAAAGTATTTCCTTATGCTAAGTGG 4355
 QY 726 ValAlaValHisIlys-----ThrValAlaAspAsnLeuValAsnProMetArgGlyTyg 743
 Db 4356 CATAGAGTGCATTAAGGGGGGAGATATCGCTGATTATATC-----CAAGAGGGCTGG 4406
 QY 744 ArgGlnArgTygSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
 Db 4407 AATGTGCTAATGGCATCCCTTATGGCAGCGAATTCGTATGAAAGAT---AATGACCT 4463
 QY 764 IleAlaArg-----AlaGlyIleSerGlyValTygSerPheGlyAspAsnAlaTyg 780
 Db 4464 ATTCCGCAAGATGTGTATATGCCAAGAAAGACGCTAATCAATATAC-----TAT 4511
 QY 781 GlySerAsnArgAlaHis---GlnMetThrGlyGlyIleGlnAlaGlyTygIleTrpSer 799
 Db 4512 GTTTTAGAGCTTACCACTTGAAGCCGACCCATGGGATCAAAAAGGGGATATGTGTAGT 4571
 QY 800 AspAsnPheAsnHisValProTyArgLeuValgPhePheAla-----Gly 814
 Db 4572 GAAGAAA-----GGCTTGTTGGCGGTGATACGTATGATAT 4607
 QY 815 GlyAspGlnSerIleArgGlyTyg---AlaHisAspSerLeuSerProIleSerAspIys 833
 Db 4608 GGTAGAGGAAGCCCGCTGCCATTATATCCCTAGCGGGTCTGTGAGATCTTGATGATGATAT 4667
 QY 834 GlyTygLeuThrGlyIleValIleValLeuAlaValGlyThrAlaGlu----- 848
 Db 4668 AGTGAAGTGAAGCACTAATATAGCGCTGATTTCTAAACCCAGGACTTAACCTTCAAAAGAT 4727
 QY 849 -----TyrAsnTyArgIlePheMetIysAspLeuArgLeuAlaValPheGly 863
 Db 4728 GCCACATGGGATCTTACAAACACCCCTTATCATTCGGAGACTTAA-----GGT 4775
 QY 864 AspIleGly 866
 Db 4776 AAGGTGGGT 4784

RESULT 30
 US-09-815-242-7313
 : Sequence 7313 Application US/09815242
 : Patent No. US20020061569A1
 GENERAL INFORMATION:


```

Db 6274 GATAGCGCGCAAGGGGTAGCGCGCGCAAAATCCGTCACTTTTCAGCGATGAGGGCTCT 6333
Oy 456 -----AspGluSerThrLeuGlu-----ProValIleGluThrValGluLeuThr 470
Db 6334 ATGACAAAGCCGCGAGCGGATCATTTGAACGCCCATTTTCTAATTTTAAACAGGCG 6393
Oy 471 AspGlyIleLeuMetCaspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAsp 490
Db 6394 TTGAATGCTTAAAGTAATCTTCAATTCACGCCATGCGCTAGAAAGCGCTTAGCGGATACA 6453
Oy 491 LysLeuAsnLeuValAlaAla-----LysAlaArgHisLeuTyrAspMetProAspAsp 508
Db 6454 GCGGCTAAACACGCAATGCGGGGATTTTGAACGCAAGAACTCATGATGTT-----TCGCA 6510
Oy 509 ArgValIleAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSer 528
Db 6511 AATGCAAGCGTGTCTGTGATGATTCGCGCACCATGAAAGGATT-----GAATCAGC 6564
Oy 529 AspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGlu----- 544
Db 6565 GAT-----ATTGGGTGGCGAGGAGCTGATTGAACTTTAGAAAGCGATTTTGGCGCG 6621
Oy 545 -----AsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThr 561
Db 6622 GTTTATTAGAAAGATGTGATGATCCCATTTGCAATGAAATCTTGTCTTTATGCGGACACT 6681
Oy 562 ProAlaAspValTyrGlnSerLysVal-----ProLeu 573
Db 6682 TTGATTTATGAAGAGGTGCTTAAAGGTGTGACCGCGGATTAATCATCATGATC 6741
Oy 574 TyrValPheValAlaSerAspLysProArgAsp-----GlyGlnIleGly 588
Db 6742 CGCACCCAGTAATCTTAAAGCCCAAGGCGCTGTCCGCAATCTTATGCTTGTAAT 6801
Oy 589 LeuGlyTrrGly----- 592
Db 6802 TTGGCGCAAGGCAAGATGATTAATCCGCTGAAGCGGTGGCTGTAGCGGCAATCT 6861
Oy 593 -----SerAspThrGlyThrArgLeuValThrLysPheGlnHis----- 605
Db 6862 ATTGGGAGCCCTGGAAAGCAGCTCATTTAAGCACTTTCCATGTGGCGGAGCAGCAGC 6921
Oy 606 -----AsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGly 618
Db 6922 AGCAGTCACGATGAGCGCAATATCGACGCAAGAAAGTTTGTGCTTTTACAAAC 6881
Oy 619 LeuArgLeuSerGlnAspLysLysGlyValLysLeuTyrAlaThrLys----- 634
Db 6982 CTTAAGACTTACAGAAATAAAGGGTAAACACATTATGCTAACCGCGTAACGCTTCT 7041
Oy 635 -----ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeu 648
Db 7042 ATTTAGTCGTAGACCTTAACATTAAGGCCCTTTGATGGGAATTTACGATTTGAACG 7101
Oy 649 GlyTyrGlnGlnGlyValPheGlnHisSerThrAsnGlyPheAspLeuSerThrArgThr 668
Db 7102 GTTTATGAAGAGTCTTCTGACG----- 7125
Oy 669 LeuGlnHisGluIleSerArgSerIleIleGlnAsnGlyLysTrrAsnArgThrTyrSer 688
Db 7126 -----GTCAAAATATGGCGATCAAGAACTAAATTTTGT 7158
Oy 689 LeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluThrTrrPlnAspLeu 708
Db 7159 TTAAAG-----ACAAGCGATATCTCAAG-----CCAAGCGAA----- 7191
Oy 709 ProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeu-----AlaGly 725
Db 7192 -----TTAGCCCGCGCTTGGCGTAAGATTCAGGCGAAAGTATTTCCCTTATGCTAGG 7248
Oy 726 ValAlaValHisLys-----ThrValAlaAspAsnLeuValAsnProMetArgGlyTyr 743
Db 7249 CATAGCGTCATTAAGGGGGAAGTATCCCTCATTTATTATC-----CAAGAGCGCTGG 7299

```

```

Oy 744 ArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
Db 7300 AATGTGCTAATCCCATCCCTTATGCGAGCGAATTTGCTAGTCAAGAT-----AATGACCT 7356
Oy 764 IleAlaArg-----AlaGlyIleSerGlyValTyrSerPheGlnAspAsnAlaTyr 780
Db 7357 ATTCGCAAGCTGTGTATGCGCAAGAAAGAAAGCGCTAATTCAAATAC-----TAT 7404
Oy 781 GlySerAsnArgAlaHis-----GlnMetThrGlyLysIleGlnAlaGlyTyrIleTrrSer 799
Db 7405 GTTTAGAGCGTAAACATTTAGACGCGACCATGGATTCAAAAAGGCGATATGGTAGT 7464
Oy 800 AspAsnPheAsnHisValProTyrArgLeuArgPheAla-----Gly 814
Db 7465 GAAAA-----GCGTTGTTCGCGTATGATGATTAAT 7500
Oy 815 GlyAspGlnSerIleArgGlyTyr-----AlaHisAspSerLeuSerProIleSerAspLys 833
Db 7501 GTTAGGGAAGCCCGCTGCCATTTATGCGTAAAGGCTTGTGATGATGATTAAT 7560
Oy 834 GlyTyrLeuThrGlyGlnValLeuAlaValGlyThrAlaGlu----- 848
Db 7561 AGTGAAGTGAAGACTAATAGCGTATTTTAAACCCAGCAGCTTAACATTTCAAAAGATT 7620
Oy 849 -----TyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGly 863
Db 7621 GCCACATGGCATCTTACACACCCTATCATTTGCGGACTTTAAA-----GCT 7668
Oy 864 AspIleGly 866
Db 7669 AAGGTGGT 7677

RESULT 31
US-09-914-300-1524
: Sequence 1524, Application US/09974300
: Patent No. US20020146721A1
: GENERAL INFORMATION:
: APPLICANT: Berk, Randy M.
: APPLICANT: Clausen, Ib Groth
: TITLE OF INVENTION: Methods For Monitoring Multiple Gene
: FILE REFERENCE: 10085, 500-US
: CURRENT APPLICATION NUMBER: US/09/974, 300
: PRIOR FILING DATE: 2001-10-05
: PRIOR APPLICATION NUMBER: 09/680, 598
: PRIOR FILING DATE: 2000-10-06
: PRIOR APPLICATION NUMBER: 60/279, 526
: NUMBER OF SEQ ID NOS: 8481
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 1524
: LENGTH: 2545
: TYPE: DNA
: ORGANISM: Bacillus licheniformis
US-09-914-300-1524

Alignment Scores:
Pred. No.: 0.00744 Length: 2545
Score: 127.00 Matches: 168
Percent Similarity: 35.72% Conservative: 161
Best Local Similarity: 18.24% Mismatches: 338
Query Match: 2.69% Indels: 255
DB: Gaps: 44

US-09-914-168-2 (1-919) x US-09-974-300-1524 (1-2545)
Oy 65 GlnIleGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeu 84
Db 260 CAGAGCAATTTGCGCTTATCTGCGAGC-----TGCTCGACGATATATGAATCTCAAGGATG 316
Oy 85 AspValAlaAsnPheAspAspGlnSerProIleSerArgIleGlyGlnGlnProPro 104

```

Db 317 AGCTTGAA-AACACGAGGCGGATGATATTCGGACCTTGCCAAATACGCGGAAAA 375
Oy LeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluGluLeu 124
Db 376 GACGGATTTCAGTGAACATTTTCCGCGCAGACAGACCTGACA-----CAGCTG 426
Oy 125 AlaGlnIleuSerThrGluMetGlyIleAsnProAsnAsp-----Tyr 138
Db 427 GCTACAGAGAGACGACGCGTCCCATCACAAGGCGCATTAACAGATGCGAATACAT 486
Oy 139 IleProGluTyrGlnGlyGluGlnProAsnSerGluValValProThrLeuGlu 158
Db 487 ACAGCTTACACATGTCGCGGAAAGATACGGA-----CTTACC 522
Oy 159 ProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsn 178
Db 523 CCTGACACAGATCATTTGATATGAAA-----GGCTGATGGGAGACTCGTCACAG 570
Oy 179 LysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyLysThrAla 198
Db 571 AATATTCAGCGGTGCGCGGA-----GTCGCGGAAAGACGCGC 609
Oy 199 IleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAsp 218
Db 610 ATC-----AACCTTTGAAACAGATTCATACAGTACAGAGAGCTCTCTCTCA 657
Oy 219 IleThrGlnIleuSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAla 238
Db 658 ATACACCAAGTACGCGGAAAAAACTGAAAGAAAAGCTTGAAGATTAAAGAACAGCT 717
Oy 239 LeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer 258
Db 718 TTGATGAGCAAA-----GAGCTTGCGACGATTAACGACAGAA 753
Oy 259 IleGlyLeuValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArg 278
Db 754 ---GCGCGCTTGAAAGTACGCTCGATTCCTCGGATAT----- 789
Oy 279 AlaValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGlu 298
Db 790 -----GAAAGTTTTCAGAGGAGGAGCGGTCGTCACAAATTTTAAAGAT 831
Oy 299 ValProLeuLeuIleGlyAspValPheHisIleGlyLysTyrGluThrLysLysAsnLeu 318
Db 832 CTCGCGC-----TTTCAT-----TCTTTG 849
Oy 319 IleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPheAsnArgSerVal 338
Db 850 CTCGAAAGGATTGGCGAAGAGCGGAGAGAAAGAAAGAACAGTTGAA---GAGATC 906
Oy 339 AspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGln 358
Db 907 GATTTTCATGATC-----AAAAACAGACATC----- 930
Oy 359 TyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThr 378
Db 931 -----ACGGATGACCTGTTTCG 948
Oy 379 AspProAspLysLeuProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMet 398
Db 949 TCACCGGCTTCCTCTCGTAGTCGAACAG-----CTC 978
Oy 399 GlyGluAlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrArg 418
Db 979 GGGCACAATTTCACGAGGCGCCGATCTCGGT----- 1011
Oy 419 TyrPheAsnMetValAsnThrGluIleValPhe-----ProGluArgGluGlnIleGln 436
Db 1012 ---TTTTCGATTCGTCACAGAACAGCGGCTTCTTCATTCGGAAGAACGCGCTTCAG 1068
Oy 437 AsnAspGlnValSerPheGlnGlnSerSerSerSerArgThrGluProAlaGlnValAsp 456
Db 1069 TCGGAC-----TCTTTTAAAGAAATGGCGCAGACATGATGCAAAAAGAAATGGGTGTTT 1122

Oy 457 GluSerThrLeuGluProVal-----IleGluThrValGluLeuThr 470
Db 1123 GATGCAAAACGTGCGCGGCTCCCTTTCGCGTGGCGGGCGCATTAAGCTGAAGCGCTGAG 1182
Oy 471 AspGlyIleLeuMetAsp-----IleSerPro-----IleGluPheSerAla 484
Db 1183 TTTGACGTTCTTCTTGCGGCGCTATATCATCAACCCGGGCGACTCGACAGATGAGGGA 1242
Oy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
Db 1243 AGCGTTGCGAAAGAACATCAATTAACATCTTTCAGCGAGAT---GAAGCGCTTACGGA 1299
Oy 505 -----MetProAspAspArgValLeuAlaIleAsnHisAspGly 518
Db 1300 AAAGAGCGGAAACAGCGGCTTCGATGAAAGAGACTTGCC---GACCAT----- 1347
Oy 519 ValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIle 538
Db 1348 -----TTAGCCGAAAGCGAAGGCGATATCGCTGCTGCGCGGAAAGCTC 1392
Oy 539 LeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsn 558
Db 1393 CTGATGAACTGGAAGAAAGCAGCAGCTCGAAGT----- 1428
Oy 559 ArgLysThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAla 578
Db 1429 -----TTGACAGCGCTCGAAATGCCCTCGCCACACATTTCTCGGT 1467
Oy 579 SerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrPglySerAspThrGlyThrArg 598
Db 1468 GAGATGAAATCGATCGTGTCCAAAGTCGATCGACAGACTGAAAAAATGGGTGAGAA 1527
Oy 599 LeuValThrLys-----PheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAla 615
Db 1528 CTGTCCGCTAAATTAGCGGAATACGAGAAAAAATCCACAGC-----TCAGCA 1575
Oy 616 GlyAlaGluLeuArgLeuSerGluAspLysLys---GlyValLysLeuTyrAlaThrLys 634
Db 1576 GGTGAACCTTCACATTAATTCACCAAAAGCAGCTCGGTGATTTGTT-----GAT 1629
Oy 635 ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluVal 654
Db 1630 AAACCTCGGGCTGCTGCTCAAAAACCAAAAGC----- 1665
Oy 655 PheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSer 674
Db 1666 ---GGCTATTCACCTTCAGCCGATGTGCTGAAAGCTGGCGGATTAACATGTCAATTATC 1722
Oy 675 ArgSerIleIle-----GlnAsnGlyGlyTyrPasnArgThrTyr----- 687
Db 1723 GAAGATATTCCTGCACTACAGGAGATGGCAAGCTCCAGTCACATACCTAGAGGCGTT 1782
Oy 688 -----SerLeuArgTyrArgLeuAspLysLeuLysThr 698
Db 1783 TTGAAGTGAATCAAAAAGACAGCCATTAAGCATATACCCTGTTAATCAACCTTAACA 1842
Oy 699 Gln-----AlaProProGluThrTyrPgnAspLeuProValAspPheVal 713
Db 1843 CAGACGGGAAGGCTCAGCTCAACGATCCGAAATCTGCAGAAATATACCGATCGCCTTGA 1902
Oy 714 AsnGlyLys-----ProSerGlnGluAla----- 721
Db 1903 GAGGCGCCGCAAAATCCGCGAGCGTTTGTCTTGCAGAAAAGGCTGCTCATTTTTCG 1962
Oy 722 -----LeuLeuAlaGlyValAlaValHisLysThrVal 732
Db 1963 GCGGATTAATCTCAGATGAGCTGAGGTTTCGCCACATTTCAAGAGATAAATTTTG 2022
Oy 733 AlaAspAsnLeuValAsnProMetLarGlyTyrArgGlnArgTyrSerLeuGluVal--- 751
Db 2023 ATTGAAGGTTTCACGAAAGCATATG---GACGTTTCACACAAAACGCGATGAGCATTTT 2079

Db	2241	TCGTGTATAAATAATTAAGAGGGAAATTTCCACAGTTATTATTCCTTCTCTACAAACA	2300
Oy	339	GlycylalanylarginineglutinalalavalaleuSerasnAspleuilealantharg	418
Db	2301	GAATGACACCTAGCAAGAAACAGAAACCACTATACAGAGATGGAACAGACTACT	2360
Oy	419	TyrPheasmetvalasnthrgluileval---PheprogluarggluInllelgnasn	437
Db	2351	TATACAGATGAAATACAAAGAGAGATCATTAAAGTCCTTTATGCGAAAAACAGACAA	2420
Oy	438	AspGlnValSerPhegluInSerSerSerSerargThrGluProAlaGlnValAspGlu	457
Db	2421	GAACCTCTCTCGGAGGAACCTCTACACTCTCTCCACGCAATTCATGTATCAGAG	2480
Oy	458	SerThrleugluProValIlegluThrValGluLeuThrAspGlyIleleuMetksple	477
Db	2481	TCT-----TCGTGGAAATATACCAG-----	2501
Oy	478	SerProIlegluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAla	497
Db	2502	-----TCCTTTGATTTCCCAACATGATN-----ACAAAGTTAGTGCA	2540
Oy	498	LysAlaargHisLeuTyrAspMetProAspAspArgValleu-----AlaIleasn	514
Db	2541	GAGCCACAGAGAGATGAGATATGAGGAGAACTTTACAGCAACTCCAGCTACTCANA	2600
Oy	515	HisAspAspGlyValAlaasnArgSerIleleu-----GlyArgIleSerAspAlaValSer	532
Db	2601	TATGATGAAATATTACAAACAGTCCTTTTGCCCATGCTACTTTAAAGTGTGAACAGCC	2660
Oy	533	AlaValAlaargAlaIleleuProAspGluSerGluAsnGluValIleAspleuProGlu	552
Db	2661	ACTGATCACAAAATGCTCATGG---GATGAAGATAAATACAACTCCAGCGCTTACAGCT	2717
Oy	553	ArgThrAlaLeuAlaasnArgLysThrProAlaAspValTyr-----	566
Db	2718	ACAGAACCTTCAGGCTCTTTCAAAATATGGCCCCCTGCTTACTCACACCTGTGGGATGAT	2777
Oy	567	---GlnSerLysLysValProLeuTyrValPheValAlaSerAspLys-----	581
Db	2778	GGAAAGATAAAGACATCCCAAGTTTCACATGAAGATGAGAGATGAATTAATTCCTTAT	2837
Oy	582	ProArgAspGlyGlnIleGlyLeuGlyTyrPglYserAspThrGlyThrArgLeuValThr	601
Db	2838	CCCAATAGTACTCAAAACAGTTAGAGAGAGTTACTGATGAAGACATGACAGCCCATGA	2897
Oy	602	LysPhe-----GluHisAsnLeuIleasnArgAspGlyTyrGlnAlaGlyValaGluLeu	619
Db	2898	AAATTCACAATTAAGATTTCAGCCCACTACTACTCACTGATATTCGCAAAAAGTCAACTTG	2957
Oy	620	ArgLeuSerGlnAspLysLysGlyVal-----LysLeuTyr	631
Db	2958	AGAAATTTCTCAACACTGAAAAAGTTCCACCTATCAACAGACATGAAGGCCAACATTTAT	3017
Oy	632	AlaThrIcyProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGln	651
Db	3018	GCAACCACTGGAAGAGAT-----GCTTTGGCTGAAGTA	3050
Oy	652	GlnGluVal-----PheGlyHisSerThrAsn	660
Db	3051	GAAGATGTGACCTCTAGCCAGATATCTAGCTTTCCCAATTTGCACACACTTCACAG	3110
Oy	661	-----GlyPheAspLeuSerThrArgThrLeuGlu-----	670
Db	3111	GTGAGAGATTAGCATTTGTTAAGTATAGAGACCCCAAGCCCTACTATTATGTAGAC	3170
Oy	671	-----HisGluIleSerArgSerIleIleGlnAsnGlyTyrPAsnArgThrTyrSer	688
Db	3171	TCCTTCCATACATTCCTCTTCTGTATATCCCAAGACAGACTGGGA-----	3218
Oy	689	LeuArgTyrArgLeuAspLysLysThrGlnAlaProGluGluThrTrpGlnAspleu	708

QY	709	ProValAspPheValAsnGlyLysProSerGlnGluAlaLeuAlaGlyValAlaVal	728
Db	3246	-----GATGAAGTTCTAGGTGAACCTCCACAGACATCTGTCTC-----ATT	3287
QY	729	HisLysThrValAlaAspAsnLeuValAsnPro	739
Db	3288	GATCAGACTCGCTTGAAGCGACTATTCTCCA	3320
RESULT	33		
US-09-954-531-1380			
	: Sequence 1380, Application US/09954531		
	: Patent No. US20020165180A1		
	GENERAL INFORMATION:		
	: APPLICANT: weaver, Joe		
	: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agen		
	: FILE OF INVENTION: Gene Sets		
	: FILE REFERENCE: 689290-77		
	: CURRENT APPLICATION NUMBER: US/09/954,531		
	: CURRENT FILING DATE: 2002-05-02		
	: PRIOR APPLICATION NUMBER: US/60/233,133		
	: PRIOR FILING DATE: 2000-09-18		
	: PRIOR APPLICATION NUMBER: US/60/234,009		
	: PRIOR FILING DATE: 2000-09-20		
	: PRIOR APPLICATION NUMBER: US/60/234,034		
	: PRIOR FILING DATE: 2000-09-20		
	: PRIOR APPLICATION NUMBER: US/60/234,509		
	: PRIOR FILING DATE: 2000-09-22		
	: PRIOR APPLICATION NUMBER: US/60/234,567		
	: PRIOR FILING DATE: 2000-09-22		
	: NUMBER OF SEQ ID NOS: 1392		
	: SOFTWARE: PatentIn version 3.0		
	: SEQ ID NO 1380		
	: LENGTH: 9534		
	: TYPE: DNA		
	: ORGANISM: Homo sapiens		
	US-09-954-531-1380		
Alignment Scores:			
Pred. No.:	0.0916	Length:	9534
Score:	124.50	Matches:	198
Percent Similarity:	34.64%	Conservative:	147
Best Local Similarity:	19.88%	Mismatches:	351
Query Match:	2.63%	Indels:	301
	9	Gaps:	50
US-09-914-168-2 (1-919) x US-09-954-531-1380 (1-9534)			
QY	12	PheMetProValAlaLeuAlaAlaIaTyLeuProLeuMet	24
Db	4470	TTTCTTCAGTACACATTTCAAGC-----CCTCTGTGTCCGACAGAGACTTGACGACT	4523
QY	25	ThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAla	44
Db	4524	ACCGGTGACGGCTTGTTCACAGGGGATATGAAGCCACGATACGTGAAGAGTGTGCCCTG	4583
QY	45	HisAspThrAlaIleAsn-----GlnAlaLysAlaGlyAsnPro-----	57
Db	4584	GCTATTACTGGCAGTCCAGGGCACCTCGAGGGCTCTCGCCAAAGATGTGATGTGATCCCT	4634
QY	58	-----ProValLeuLeuThrProGlnGlnIleGlnAlaArgLeuAsn	71
Db	4644	ATGGCTACTGCTGCTGCTGCTGTGACCCCTGTACAGAGATTTCACAGTGTCCGACCTGGAG	4703
QY	72	-----AlaAlaGlyLeuAsnAlaLysProGlnSerGln	82
Db	4704	CCAGGGAAGGAAGTGTGACGGCTGCAGACGACATCGCATCGACGGCAGG-----	4751
QY	83	AlaLeuAspValValAsnPheAspAspGlnSerProIleSerArgIleGlyGlnGlnSer	102
Db	4751	-----	4751

Oy	103	ProProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluGlu	122
Db	4752	-----CGTGGAGTGTGTGTTTTGTGGG-AATAGTGGACGTGGCTTCTGTGGTGAC	4804
Oy	123	LeuPheAlaGlnGluSerThrGlnMetGlyIleAsn-----ProAsnAsp	137
Db	4805	TTGGCTCGCGCTGGACACAGATGGTATGACATCACTACTACTGTGCTCCGTGGCGCA	4866
Oy	138	ThrIleProGluThrGlnGlyGluGlnProAsnSerGluVal-----ValValProPro	155
Db	4865	TATAAATGCTGTATGAGCTTGAAAAATGATGACACAGAGCTAAAGCACTTGCTGTACCT	4924
Oy	156	ThrLeuGluProGlyLeuProGlyLeuIleValArgGluGluThrIleArgGluPheAsnAsp	175
Db	4925	CACGGGCCCCACAGAGG-----CTTATTTCAGCTGGCACAGGCGAATGTG--AATACA	4975
Oy	176	GlyValAsnIleValProArgLeuIleValAlaLysPheThrGlnSerSerGlnSerGluGlu	195
Db	4976	CTCGTACCGCAAAATGACAGCACTGCTGACAGGGATCCAAAGTGCACAGATGGCGAG	5035
Oy	196	ThrSerAlaIleGlyIleSerSerHisGlnIleThrGluProGlyIleAsnIleLysAlaIle	215
Db	5036	CAGACC-----GGACAGATGCTGGAAGGACACACACAGAGCAAG-----TCC	5080
Oy	216	LeuGluAspIleThrGlnGlnGluSerAlaMetArgP-----LeuAsnGlySerIlePro	232
Db	5081	CTGGGAGCAATTCATTAAAGACCTTGCCCGGATGCGAAGAGCTGTAATGAAAAACTATA	5140
Oy	233	ArgLeuArgGlnThrIleAlaLeuValAlaAlaArgAlaValGlyIleThrIleAspIleAsp	252
Db	5141	AAACTAAATGCAACT-----CTAGCACTGCA-----GACAG	5173
Oy	253	SerIleIleArgAsnSerIleGly-----GluValAspValIleIleHisAspLeu	269
Db	5174	GCCTTTGAGCAAAATTTGGAAAGGCTTCACAAAGAGATTGGCCAGATGTTAAAGAACTG	5233
Oy	270	GlyGluProValThrIleAspThrArgAlaValGluValArgGlyGluGlyValAspAsp	289
Db	5234	AGGAGC-----AAAAATCTAGACACACAAAGCAAAATTTGCTGCATGAT	5275
Oy	290	Lys-----AlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPhe	307
Db	5276	GAGTTGGTCACTGCAGAAAGCCCTTGCAAAAAGTGAAGACACTGTTGGAGAGTCCGG	5335
Oy	308	HisHisGlyIleGlyThrGluThrIleLysIleAsnLeuIleGluAsnAlaSerAlaGluHisGly	327
Db	5336	GGGCAAAATGAA---GAATGGAAGAGATCTCCGGGAAAACCTGGCTACCAAAAAC	5392
Oy	328	ThrPheAspGluArgThr-----LeuAspArgSerValAspValIleLeuProAspAsn	345
Db	5393	AAAGTTGATGTCGCTGGGACCTTTTGAGAGAAAGCCACAGATAAATAAGCAAGCTAAT	5452
Oy	346	ThrAlaAspValSerLeuIleThrIleAspThrGlnThrClnThrArgPheAspGluValVal	365
Db	5453	CGC-----	5455
Oy	366	PhePheThrIleAspProLysThrAsnGlnLeuThrIleAspProAspLysLeuProVal	385
Db	5456	CTATTTCAGCAATTAATCGAAA-----ACATGACTGCACTTTGGAGCAAA	5497
Oy	386	LysArgGluLeuLeuGlnGluGlnLeuLeuThrValAsnMetGlyGluAlaThrAsnLeuGln	405
Db	5498	AAGAGAGACGCTGTGGAGAC-----GGCAAAACA---CAAAATTGAG	5536
Oy	406	AlaValAlaGluLeuSerAsnAspLeuIle-----AlaThrArgThrPheAsnMetVal	423
Db	5537	AACACTTTAAAGAGCAATGACATCAATCTCGATGAAGCCACACCGTTCGACGATGAATC	5596
Oy	424	AsnThrGluIleValPheProGluThrArgGluGlnIleGln-----	436
Db	5597	AATCTCATCTATACAGTATGTTGAAGCAATCCAAACTAAATTGCCACCTATGTCTGAGAG	5656
Oy	437	-----AsnAspClnValSer-----PheGluGln	444

[illegible]

Dh 6581 GTTGCCTAATCACTCTC-----TTTTATCT 6607
Qy 752 glyserSerGlyLeuValSerAspAlaHisMetAlaIleAlaArgAlaGlyIleSerGly 771
Dh 6608 GGAAGTGCACCAATTATTGACTTCTGCTATAGAAATGCGTAAGGCAAGTCAAGCTTC 6667
Qy 772 ValTyrSerPheGly-----AspAsn 778
Dh 6668 CTCGGGATGTTGGATCTGGAGTTGGAGCTGAGTACCCAGATTGACTATTATGAC 6727
Qy 779 AlaTyrGlySerAsnAlaGlyAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTrp 798
Dh 6728 TCATTTGTGATCCGTATCGTAGCATCAAGAACTGG-----AGAAAGAACTATT--- 6778
Qy 799 SerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspSer 818
Dh 6779 -----TCGTGAGAGCCCTGGATGAGCCCAAGCCAGC 6811
Qy 819 IleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspIysGlyTyr----- 835
Dh 6812 ATTTGCCACACACACACATTCGACGCTCTCT-----CCAGGATACAGATTCTA 6862
Qy 836 ---LeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPheMet 854
Dh 6863 GATGTGATGCAATGCAATGCTGTTGTGTGGCTGACTGGGAATTAAAGAAAGCT 6922
Qy 855 LysAspLeuArgLeuAlaValPhe---GlyAspIleGlyAsnAlaTyr 869
Dh 6923 GATCTGTAGCTGTGATTACATTCTAGCTGCTGATGGAGAAACATAC 6970
RESULT 34
US-09-782-378A-23
: Sequence 23, Application US/09782378A
: Patent No. US20020102731A1
: GENERAL INFORMATION:
: APPLICANT: Hearing, Patrick
: APPLICANT: Bahou, Nadie
: APPLICANT: Sandalon, Ziv
: APPLICANT: Gatenko, Dmitri
: TITLE OF INVENTION: Adenoviral Vectors
: FILE REFERENCE: STONYB-04970
: CURRENT APPLICATION NUMBER: US/09/782,378A
: CURRENT FILING DATE: 2001-02-12
: PRIOR APPLICATION NUMBER: 60/223,747
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 27
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 23
: LENGTH: 10302
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-782-378A-23
Alignment Scores:
Pred. No.: 0.103 Length: 10302
Score: 124.50 Matches: 138
Percent Similarity: 37.35% Conservative: 119
Best Local Similarity: 20.06% Mismatches: 258
Query Match: 2.63% Indels: 173
Dh: 10 Gaps: 36
US-09-914-168-2 (1-919) x US-09-782-378A-23 (1-10302)
Qy 146 GlnProAsnSerGluValValProProThrLeuGlnProGluLysProGlyLeuIle 165
Dh 5440 CAAAGAGAGAGAAATGTTTACATCAACCTTATGAGAGATTAATAAAAAAGATC--- 5496
Qy 166 LysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAla 185
Dh 5497 ---CGTTTGCAATATTACTTTTGCATCTAGATCAACAAAT---AAGCA 5544
Qy 186 LysPheTyrGlnSerGlnSerGlnGlu---ThrSerAlaIleGlySerSerHisGln 204

Dh 5545 ATCCCTATTCAACAGAGAAATGGTCAACTGCTTCTGATTAATGATCATCACTTCT 5604
Qy 205 LysThrGluProTyrAlaAsnLleLysAlaIleLeu-----GlnAspIleThrGln 221
Dh 5605 CCTACAGATTATCTGGTTGAATTAACAAATTTTACTTTTGCATGGATGATGTT----- 5658
Qy 222 GluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAla 241
Dh 5659 -----GAATTATCGCTTAATGTTCCAGAGCT---AACACTGCTATTTC--- 5700
Qy 242 AlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArg-----AsnSer 258
Dh 5701 ---GAAGACTTCTCTTTTCAGAGAACCTCTGAGAAATATACAAAGCCAACTGGACAA 5757
Qy 259 IleGlyGlnValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyr--- 277
Dh 5758 CTYGGAGAGCAGATTGCATGCTCATGTGAAAAACAGCCAGATGTCATCTTGAAGCTCT 5817
Qy 278 -----ArgAlaValAlaGluValArgGly----- 284
Dh 5818 GGACCTGAAGCCATTTCAGATCAGATACACTTACTCAGCTGAATGCAAAATGGACAGA 5877
Qy 285 -----GluGlyAlaAspAspLysAla----- 291
Dh 5878 ATTAATAGATGTACAGATGATCGAAAGGTTGTTTTCAGAGGCAATGGAAGATTGAGA 5937
Qy 292 -----GluGlyAlaAspAspLysAla----- 291
Dh 5938 CAGTTCCATTGTGACCTTAATGACCTCACACAGTGGATTAACAGAGCTGAAGAA----- 5991
Qy 301 LeuLeuIleGlyAspValPheHisAspGlyLysTyrGlyLeuThrLysLysAsnLeuIleGlu 320
Dh 5992 TTACTGGTTGATACCTGCTGCTCCAGCTGACCTGATTTAGAGAAACCCAGATACAT 6051
Qy 321 AsnAlaSerAlaGluHisGlyTyrPheAspGlyArg-----TrpLeuAspArg 336
Dh 6052 CAGCAGAACTTGAGGTGGCTGATCAGACCCAGCCAGCTTTTTCAGACCACTTAACCGA 6111
Qy 337 SerValAspValIleLeuProAsp---AsnThrAlaAspValSerLeuIleTyrAsp--- 354
Dh 6112 ACTGGGATGGATTTGTGAGAACTCTCCAGGCAAGATGGAAGCTTCTTGAAGAGAAAA 6171
Qy 355 ---ThrGlyThrGlnTyrArgPheAspGlyValValPhePheThrIleAspProLysThr 373
Dh 6172 CTGGCAGGTTTAACCAACGCTGCGATGCAATTGTT---GCAGAGTGAAGATGAGCAG 6228
Qy 374 AsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeuGlnLeu 393
Dh 6229 CCAAGGCTAAAGAGAAAGTAAGCAGGTGATGAAGTACAGCATCAGCTATGATGAGAT 6288
Qy 394 Leu-----ThrValAspMetGly 399
Dh 6289 ATCTGTTGTTAACAAGGCTGACATGCTATGCAAAAGATCAACCCCAATTGGGA 6348
Qy 400 GlnAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAla----- 416
Dh 6349 GAA-----AACCTCAGAAATTAAGAGACTTAACCAAGAAATGAAGTACATGCTGAA 6402
Qy 417 ---ThrArgTyrPheAsnMetValAsnThrGluIleVal-----Phe 429
Dh 6403 AAACCTCAATGCTGTAAGTAAGTAAGTGAATGATGATCTTCAGATAAAGTCTGATTGA 6462
Qy 430 ProGluArgGluGlnIleGlnAsnAsp-----GlnValSerPheGluGlnSerSer 447
Dh 6463 CTTGAAGGATAAATTTTCAGAAAGCTTAAGGACTGTAAATATGACATGATTAAGATT 6522
Qy 448 SerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrVal 467
Dh 6523 TGCAGAGAGGTGGCTACACACCTTGAAGATCATCCAGAGACC----- 6567
Qy 468 GluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeu 487
Dh 6568 -----AGTTCTGTTTCACAGACAGAGATGCTGCTATCTTAAT 6606

```

OY 488 ILeGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetPro--- 506
Db 6607 GTCCAA---AAGGTGCTGCTAGTATCATCTCGCA-----GATATTCCTGTT 6651
OY 507 ---AspAspArgValLeuAlaAlaLeuAsnHisAspAspGlyValAsnArgSerIleLeuGly 525
Db 6652 CAGTCTCATCTGACTTCGGAATTTCAATTCCTGCTGATCTTGATTAACATTAACA--- 6708
OY 526 ArgIleSerAsp-----AlaValSerAlaValAlaArgAlaIleLeuProAspGlu 542
Db 6709 GAAGTACCCGACTGCTGCTGATTAATGACACAGATGCTGAAGTCCAACTTGCTACTGTT 6768
OY 543 SerGlnAsnGlnValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrPro 562
Db 6769 GCGGATCTAGACAGATCAAT-----AAGACCTTTCCCGAATGAAATTAACAAG 6819
OY 563 AlaAspValTyrGlnSerLysValProLeuTyrPheValAlaSerAspLysPro 582
Db 6820 GCTGACTTAGACACAGCCCTCCTCAGCTGATTAATGTTTTCATTTGCGACAGAAATTGG 6879
OY 583 ArgAspCysGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuValThrLys 602
Db 6880 AAAAAT-----AAAGCTTCAGTTCCAGATATGAAACACCATTTACAGAAA 6927
OY 603 PheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlnLeuArgLeuSer 622
Db 6928 TTGGAACCGCTCAAGAACCGATGGATGGACCCACCATGCGCTGAGCTAAGACAGCAG 6987
OY 623 -----GluAsp----- 624
Db 6988 CAGCTTAGACAGACATGATTTGACAGCTTTGAGTGGATAGGAGAGAGACTGAA 7047
OY 625 -----LysLysGlyValLysLeuTyrAlaThrLys-----Pro 635
Db 7048 GAAGTATGAGAAATATATGAGGCTGAGCTGATATTTCCACCAAGCCGACGGATCCA 7107
OY 636 LeuSerHisProLeuAsnAspGln-----LeuArgAlaThrLeuGlyTyrGlnGlnGlu 653
Db 7108 CTCACCAACAAATTTCTGATTAACCAATACTGCTTCAAGAACTGGGCTCGAGATGCT 7167
OY 654 ValPheGlnHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuHisGlnIle 673
Db 7168 ATC-----GTCATGGCGTTCAT-----AACGTC 7191
OY 674 SerArgSerIleIleGlnAsnGlyLysTrpAsnArgThrTyrSerLeuArgTyrArgLeu 693
Db 7192 CTGCGAAGACCTCGAGAGATATGAGGAGTATGACACAAAGATGTGAAGAAACACACA 7251
OY 694 AspLysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheVal 713
Db 7252 GAGTACTTAAAAACA-----TCATGATCAATCTCAAAACAAATATTGCT 7296
OY 714 AsnGlyLysProSerGlnGluAla 721
Db 7297 GACAGACAGAACCGCTTGAGGCT 7320

RESULT 35
US-10-108-605-102
; Sequence 102, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; FILE REFERENCE: PROTEINS ESSENTIAL FOR LARVAL VIABILITY AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16

```

```

; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 102
; LENGTH: 14155
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-10-108-605-102

Alignment Scores:
Pred. No.: 0.163
Score: 124.50
Percent Similarity: 33.04%
Best Local Similarity: 20.00%
Query Match: 2.63%
DB: 9
Gaps: 41

US-09-914-168-2 (1-919) x US-10-108-605-102 (1-14155)

OY 110 SerValIleGlnGluThrThrProLeuSerLeuGluLeuPheAlaGlnGlnSerThr 129
Db 8934 AGTGCCTGGAGACTTAACAGGTAAACCTGGCTTTTCCACCAAGTCAAAATTTGAA 8993
OY 130 GluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlnGlnGlnProAsnSer 149
Db 8994 TATTACGACCCAGCTGGCCGATGAACCTGAGCCCAAGGTTAACTATTGATCCCAACAC 9053
OY 150 GluValValValProProThrLeuGluProGluLysProGlyLeuLysArgLeuTyr 169
Db 9054 GTGACCTGAGTCCATCC-----AAAAAGCCCAAC 9083
OY 170 AlaArgLeuPheAsnAspGly-----ValAsnLysValProArgLeuLysAlaLys 186
Db 9084 AGCGAATTTGAATTCGATGCGTCAAGTCCATATGCCAAGACGTCACACACACCTGGCAAC 9143
OY 187 PheTyrGlnSerSerGlnSerGlnSerGlnGlnThrSer----- 197
Db 9144 GCCCTGATATCCGTGAGGATGAGACACCTTTGGTAAACATTAACAGTTCCTTATGAT 9203
OY 198 ---AlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLys----- 213
Db 9204 GAGCGGCTCAAGAGTGGCCGACGACGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 9263
OY 214 -----AlaAlaLeuGlnAspIleThrGln 221
Db 9264 AAGAAATCTGAGCGGAGCGAGACGACAAAGATGATGCTGCTGAGACAA---GCACAG 9320
OY 222 GluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAla 241
Db 9321 CACATTTTGGGCAAAATACGACGACCGATGCAACTTACACCCANTGAACAAAGTCTG 9380
OY 242 AlaArgAlaValGlyTyrTyrAsp----- 249
Db 9381 GAAAGGCGAGGAAACCTATGACAAAGTGAACCTAGTTCGCCATCAAGCGCA 9440
OY 250 ---IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAsp 268
Db 9441 AACAGAGACCTCAATGCCCTCAAGAAATGATTTGAGAAATTCAGCGACCATCTGGAAGAT 9500
OY 269 Leu-----GlyLysProValTyrIleAspTyrArgAlaValGluValArgGly--- 284
Db 9501 CTTTCAACTGGAGCGAAGCTTCACAGCCCAAGTCCGCTGACCGCCGCAATGTG 9560
OY 285 -----GluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeu 301
Db 9561 GCCAAGCAGAGGCTTTGACAACTCAATTCGACACCGTTTCGAGCA--- 9611
OY 302 LeuIleGlyAspValPheHisHisGlyLysTyrGlnThrLysLysAsnLeuLeuGln 321
Db 9612 -----AAGCTACAGCGAGAGAGAAC----- 9632
OY 322 AlaSerAlaGlnHisGlyTyrPheAspGlyArgTyrLeuAspArgSerValAspValIle 341

```

```

Db 9632 ----- 9632
Qy 342 LeuProaspasnThrAlaaspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPhe 361
Db 9633 ----- 9650
Qy 362 AspGluValAlaPhePheThrIle-----AspProLysThrAsnGlnLeuThrThrAsp 379
Db 9651 ----- 9695
Qy 380 ProAspLysLeuProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMetGly 399
Db 9696 TTGGATAACTTG-----AGAGATGCTCTAAACGAGCTCAATTCGTCATCAATAA--- 9743
Qy 400 GluAlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrArgTyr 419
Db 9743 ----- 9743
Qy 420 PheAsnMetValAsnThrGluIleValPheProGluArgGlnGln----- 434
Db 9744 ----- 9791
Qy 435 ----- 9792
Qy 451 GluProAlaGlnValAspGluSerThr-----LeuGluProValIleGluThrValGlu 468
Db 9852 CTGGCTCCCAATACACGGATGATGCTCCAGTCCGCGCGGCGCATCAACGCCCAACT 9911
Qy 469 LeuThrAspGlyIleLeu-----MetAspIleSerProIleGluPhe 482
Db 9912 GCCTACCTCGGAATCGTTGAGCGCGTTGACCTGCCCGAAGACATCAGTCAGATGCAATC 9971
Qy 483 SerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaValAlaArgHisLeu 502
Db 9972 TCTCGCGCGGAAATGACCTACTGATAGACATGATGCAATGGAAGAGAGCT---CATCTC 10028
Qy 503 TyrAsp-----MetPro 506
Db 10029 GCTGACACCTGATCTACTGATCTTCTCCAGACACCGCCAGTCTCTGACAGAGGTCAA 10088
Qy 507 AspAspArgValLeuAlaIleAsnHisAspArgly-----Val 519
Db 10089 GAGACACCTAGAGCCCGCTGACCGCTGACCGCGGCAAGGTCGCAAAATCTCACCCCTG 10148
Qy 520 AsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeu 539
Db 10149 AACAAATCCACCGAACACACAGCTGATTAACAAGCTGATTAACCA-----CTG 10202
Qy 540 ProAspGluSerGluAsnGlnValIleAspLeuProGluArgThrAlaLeuAlaAsnArg 559
Db 10203 CCAGCTAGTCCGAGAGGATATG-----TGGAAATTTCAATGCCAAT--- 10247
Qy 560 LysThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPhe-----Val 577
Db 10248 -----GCCAGTGAATGCTCTGAGATTAAGTGAAGAACCTCTAAGATTAAGTGAAGCTG 10301
Qy 578 AlaSerAspLysProAlaGspGlyGlnIleGlyLeuGlyTyrPglYSerAspThrGlyThr 597
Db 10302 AGTGTCCAAACGCCAAAGGACCTGACAGAACGACATGGCATAGAGATTG----- 10355
Qy 598 ArgLeuValThrLysPheGlnHisAsnLeuLeasnArgAspGlyTyrGlnAlaGlyAla 617
Db 10356 -----GATTCGACCAATAGAGCGTTTCTCAACGCTACAA 10391
Qy 618 GluLeuArgLeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSer 637
Db 10392 CAATTG-----GATGACGTAAGAGCA---TCCGTTTGAAGCTAAAGCAATTTGCA 10439
Qy 638 HisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGlnValPheGlyHis 657

```

```

Db 10440 GAGGATATCGAGGAACAG---CAGCACCGTGGGCGAGCTACAGAGCCGAGCTGGGCCAG 10496
Qy 658 SerThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHisGlnIleSerArgSerIle 677
Db 10497 GAATTCGAAACTTGAAGGCCCACTGAGGCT---GCTCGGCACTTGGCCAAACGATC 10553
Qy 678 IleGlnAsnGlyLysTyrAsnArgThrLysSerLeuArgTyrArgLeu---AspLysLeu 696
Db 10554 AAGGGGGCGCTCAATTTCAACCGCAAGCAGATCTCGGAACGTAAGACACCGGAAAGACC 10613
Qy 697 LysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLys 716
Db 10614 AAGTTGCTAGCC-----ACTGCAACATATATGACCTATTTCGGCACCGCAG 10664
Qy 717 ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsn--- 735
Db 10665 CCATCCGTTTCTTTTATCTGCGC---AATGCAACACAAACGCCGCAAGAAAC 10721
Qy 736 -----LeuValAsnProMetArgGlyTyrArg---GlnArgTyrSerLeuGlnVal 751
Db 10722 GACTTTGCGCCGTTGAGATTTGTAATGCTATCCGATTTCCATAGATTGGGCAAT 10781
Qy 752 GlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyLysSerGly 771
Db 10782 GGGCGGAACGCAATCACCAGCGATGATGACGATGCGCATGCTGATATCACTGCTTT 10841
Qy 772 ValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGly 791
Db 10842 GTGATCGCATGATGCGCAATGCTCAAGCTAACATATAGGAAGAACTAACCCACGGTAC 10901
Qy 792 Ile-----GlnAlaGlyTyrIle---TrpSerAspAsnPheAsnHisValPro 806
Db 10902 GTGCTTGACAGCAGCAAGTCTGCTATCTTGAGAGATCCCAATATCTTCATGCTGAT 10961
Qy 807 TyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyr---AlaHisAsp 825
Db 10962 AAGAACACTCCCTATTTGTTGTTGCTATCCGGGTATTTGCGATTCATCAATCTCCGCA 11021
Qy 826 SerLeuSerProIleSerAspLysGlyTyrLeuThrGlnGlnValLeuAlaValGly 845
Db 11022 GACTTGACCACTACTCTTCTCCGCGACATGAGATCTTAACATTTGGCGATGACAC 11081
Qy 846 ThrAlaGluTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIle 865
Db 11082 GTAGATTTGTGAACCTT-----GTATATGCCAT--- 11111
Qy 866 GlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAla-----Gly 882
Db 11112 -----GATAACGACCAAGGCGCGCCAGGAGCGGAT 11141
Qy 883 ValGlyValArgTyrPheLysProValGlyGlnValArgValAspValAlaThrGlyVal 902
Db 11142 GTGCTGCTCGAGACAGAACAGCCGGTA-----ACTGCTG 11177
Qy 903 LysGlnGlnGlyAsn 907
Db 11178 CGGTTCAGAGGTAAAT 11192

RESULT 36
US-09-808-880-1
; Sequence 1, Application US/09808880
; Publication No. US20030027287A1
; GENERAL INFORMATION:
; APPLICANT: Belach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/808,880
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US/09/428,517
; PRIOR FILING DATE: 1999-10-28

```



```

Db 22602 AGTGGCGTCCGGCGACATTCTTTCTTGACGAGCTGCACGAGCCACCGCCGCCGCCCTTG 22661
Oy 574 TTTValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrPheGlySer 593
Db 22662 TTCACGGCTGGCAGCAGACACG-----GGACAGCCCGCGGAGCAG 22700
Oy 594 AspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArgAspGlyTyr 613
Db 22701 GACGGCGGTAGCGACACTGCTACGAGCTGCG----- 22730
Oy 614 GlnAlaGlyAlaGlnLeuArgLeuSerGlnAspLysLysGlyValLysLeuTyrAlaThr 633
Db 22730 ----- 22730
Oy 634 LysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGlu 653
Db 22731 CTTCCGCTACCGAG-----ACGGAGCAACAGCAGAGCTGCTGCGCTCGTCCAGACT 22784
Oy 654 -----ValPheGlyHisSerThrAsn----- 660
Db 22785 GAATGCTGTCGCTCTAGCGCACCTCTCCACCGAGCGGTCACAGCCAGCAGCGCGCTTC 22844
Oy 661 -----GlyPheAspLeuSerThrArgThrLeuGlnHisGlnIleSerArgSerIle 677
Db 22845 CGAGAGATCGGGTCTGACACTGACAGCCGCTC----- 22877
Oy 678 IleGlnAsnGlyGlyTyrPasnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLys 697
Db 22878 -----CAGCTCCGG 22886
Oy 698 ThrGlnAlaProGlnThrThrTrpGlnAspLeuProValAspPheValAsnGlyLysPro 717
Db 22887 AACCGGCTTACGGCCACACGCGGATCGCTCCGACAAAGCTGCTCTTCGACTACCGG 22946
Oy 718 SerGlnGlnAlaLeu-----LeuAlaGlyValAla-ValHisLys 730
Db 22947 ACCACCAACGAGACTCGCCGAGTACCTGCTCCGAACTGTTCGGTCTGCGCGGACCA 23006
Oy 730 sThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuG 750
Db 23007 G-----CTACCTCTCCGTCGTCGTCGCGAACCAGCGGATGAGAGAGACGACGCC 23051
Oy 750 uValIlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSe 770
Db 23052 GTGCGATCGTGGGATGCGCTGCGCTGCCGCTCCGCGCGGATGATGACCGGAAAGCTTC 23111
Oy 770 rGlyValTyrSer-PheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrG 790
Db 23112 TGGAACTGCTCGAAGCGG-----CGGCGATGCTCATCTCCGAACCTT 23153
Oy 790 LysGlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArg 808
Db 23154 CCGGCA-ACCGCGCTGGGACATGAGGACTCTCGAACCAGCGGAGCGCA 23208

RESULT 37
US-10-107-649-1
: Sequence 1, Application US/10107649
: Publication No. US20030044949A1
GENERAL INFORMATION:
APPLICANT: Behr, Regine
APPLICANT: Sloma, Alan
TITLE OF INVENTION: Polypeptides Having Gamma-Glutamyl
FILE REFERENCE: 10157,200-US
CURRENT APPLICATION NUMBER: US/10/107,649
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: 60/279,374
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1815
TYPE: DNA

```

```

: ORGANISM: Bacillus agardhaerens
US-10-107-649-1

Alignment Scores:
Pred. No.: 0.00908
Score: 124.00
Percent Similarity: 37.12%
Best Local Similarity: 20.54%
Query Match: 2.62%
DB: 9
Gaps: 30

US-09-914-168-2 (1-919) x US-10-107-649-1 (1-1815)

Oy 190 SerSerGlnSerGlyLuhThrSerAlaIleGlySerHisGlnLysThrGluProTyr 209
Db 163 AGCGAAATTAATAGTCAACCGCTGACAAATTAATAGTGAAGGCTGCTGATATTAT 222
Oy 210 AlaAsnIleLysAlaLeuGlnAspIleThrGlnGlySerAlaMetAspLeu----- 227
Db 223 GGCACAGAGCTCCGCC-----CATCCGTTAGCTGCGAAGATTGGAATGACATTATCGAG 276
Oy 228 AsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyr 247
Db 277 AATGGCGGT-----ACTGCTATGATGCCCTGCTGCTGCTTCATT 318
Oy 248 TyrAspIleAspLeuSerIleIleArg-----AsnSerIleGlyGluValAspVal 264
Db 319 ATC-----CTTACCTAGTTGAGCCTTACGGATTCGGATTGGAGCGGTGGGCTT 369
Oy 265 IleIle--HisAspLeuGlyGluProValTyr--IleAspTyrArg--AlaValGlu 281
Db 370 ATGCTTACACAGCAGCAGACAGAGGGTCAATAGTTATGATATATCGTGAAGCGCCGT 429
Oy 282 ValArgGlyGluGlyAlaAspAsp----- 289
Db 430 ATAAGCGGAAT-----ATAGCCCTACTGCGGAGGGCTTCAATCCCTGGCTTGT 483
Oy 290 LysAlaPheThrThrValAlaAsp----- 297
Db 484 AAGGAAATGACCTCATTTATGACATCATGGAAGAACTCCGCGGAGACGTTATCGCC 543
Oy 298 -----GluValIProLeuLeuIleGlyAspValPheHisHisGlyLys 311
Db 544 CTGCGCATTAAGACGGGCTGAACACAGCTTTTACAGTAGTGATATTTTCCACGAAACA 603
Oy 312 TyrGluThrLysLysAsnLeu--IleGlnAsnAlaSerAlaGlnIleGlyTyrPheAsp 330
Db 604 GGTAAATGCTTTCCTTACCTTGAATGAAAGAAACAGACAGCGTCACTATTCTTCCGAA 663
Oy 331 GlyArgTrpLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSer 350
Db 664 GAGACAGCTCTCGGTATGACCACTGCTTCAAGAAAGATTTTACCAGATACGTTACGA 723
Oy 351 LeuIleTyrAspThrGlyThr----- 357
Db 724 CTCATTCAGAGAGAACCGGTCGATGTTTACAGTGACCAATCGGGATCTTTACAA 783
Oy 358 ----GlnTyrArgPheAspGlyValAlaPhePheThrIleAspProLysThrAsnGlnLeu 376
Db 784 CAACAATTCAACTTACTTAAGAGATCTGCGTATGTAACACAAATTAACGTAA--- 840
Oy 377 ThrThrAspProAspLysLeuProValLysArgGlnLeuGlnGlnLeuThrVal 396
Db 841 -----CTGTCTCTCGTGAAGTGGAGACAAATTTGCTATGCT 879
Oy 397 AsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAla 416
Db 880 GGTCCCTCACCTTCATCTGAGACAGTACGTCGTCGCAAGCTTGG----- 921
Oy 417 ThrArgTyrPheAsnMetValAsnThrGlnIleValAlaPheProGluArgGlnIleGln 436
Db 922 ---CAAGTGGCAGATCAGTCGATTAAATGAGAGTTTCCCGGAT---GAAAGACGTGCC 975

```



```

OY 437 AsnAspGlnValSerPheGluGlnSerSerSerArgThrGluProAlaGln----- 454
Db 976 GAAGAT-----TTCTCATTCGAGACCTTGTCTAACTGTCAGATTCACAGACATT 1026
OY 455 -----ValAspGlnSerThrLeuGluProValIleGluThrVal 467
Db 1027 TATATTCATTGTAATGAATGAATAACAAAGCCACCTATGACGCCCTTGACACT--- 1083
OY 468 GluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeu 487
Db 1084 ---TTAGCTGACCCCTGCCCTTGATGATATTCATCATCAAGCACTACAGATGATCATTTAT 1140
OY 488 IleGlnAspLysLeuAsn-----LeuValAlaAlaLysAlaArgHis 501
Db 1141 ATTCAAACATTACTAGACGATATTTCTTTATATGAGATTACCTGGTATACCTCCGAG 1200
OY 502 LeuThrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArg 521
Db 1201 TTATTCGATTTCTCTCGCAGAGAAAGCCGATTTCTGACATACCACCTATTCGTTATGTGA 1260
OY 522 SerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArg-----Ala 537
Db 1261 GATAAAGAGCAGCAATGCTATCCGCTACCATTCACCTAGTGAATTTTGGTTCTGCT 1320
OY 538 IleLeuProAspGlu-----SerGluAsnGluValIleAspLeuProGluArgThrAla 555
Db 1321 ATTTATATTGACGGCTTTTATATCAATATCAATGACTAATTTTATGATATATCCACAT 1380
OY 556 LeuAlaAsnArgLysThrProAlaAspValIleGlnSerLysValProLeuThrVal 575
Db 1381 TCATTTATGCTTTATGAGCCGAGTAAAGGCCAGTACTTTTTCGCCACCATG---ATT 1437
OY 576 PheValAlaSerAspLysProArgAspGlyIleGlyLeuGlyThrProGlySerAspThr 595
Db 1438 TTT-----GAAGAAGAAAGCCAGCTTCGGGATGCGGTCCACCAAGT 1482
OY 596 GlyThrArgLeu-----ValThrLysPheGluHisAsnLeuIle 608
Db 1483 GGAAGACGATATCTCTGCTATGCTATTTCACAGCATTCATGCATATAT---CATTAAGAATA 1539
OY 609 AsnArgAspGly-----TyrGlnAlaGlyAlaGluLeuArgLeuSerGluAsp 624
Db 1540 AATGACGATGCTGATCCATGACACTTCACAGAACGCATCGAGGCTCCCGCTTTTATATAC 1599
OY 625 LysLysGlyValIleLysLeuTyrAlaThrLysProLeuSerHisLysProLeuAsnAspGlnLeu 644
Db 1600 GAAGAAAGATGCT---ATCTATTACAAAGAGGAATTAACCTGAAGATGTTAGCAAGAACTT 1656
OY 645 ArgAlaThrLeuGlyTyrGlnGluValPheGlyHisSerThr 659
Db 1657 CCC---AATATGGGATAC-----TCGTGTTGTGACATAGCTCA 1692

```

```

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 5346
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-738-877-2

Alignment Scores:
Pred. No.: 0.0441 Length: 5346
Score: 124.00 Matches: 180
Percent Similarity: 34.07% Conservative: 130
Best Local Similarity: 19.78% Mismatches: 301
Query Match: 2.62% Indels: 301
DB: Caps: 45

US-09-914-168-2 (1-919) x US-09-738-877-2 (1-5346)
OY 31 GlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHis----- 45
Db 2825 CAGAAAGAAAGAACCCCGCCGCTTACTGAACCT-CTGCCAGAAACAGAGAGCCCGGGGC 2883
OY 46 AspThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGluGln 65
Db 2884 GACAGGGTGTAGTGAAGCGGAA-----TTGACCCCGAAAGCT 2922
OY 66 IleGlnAlaArgLeuAsnAlaIleGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAsp 85
Db 2923 GTGACAGCTGCAGAAACCTCCAGGGCCATTTGGTTCCGAAAGAGAACCGAAGCATCTCCT 2982
OY 86 ValValAsnProAspAspGln---SerProIleSerArgIleGlyGluGlnSerProPro 104
Db 2983 GCTGAAGAGACCACAGAAATGTTGTCAAGACTCTCCAGTTAACCGAC---TCCCCA--- 3036
OY 105 LeuGlyLeuAspMetSerValIleGluGlnThrThrProLeu----- 118
Db 3037 -----GACACCACAGAGAGGCCACTCCGGTGCAGAGGTGGAAGGTGCC 3081
OY 119 -----SerLeuGluGln 122
Db 3082 GTACCTGACATAGAAGAGCAAGAGAGCGGAGCTCAAGAGTCTCCAGGACGTGGCAGAA 3141
OY 123 LeuPheAlaGlnGlnSerThrGluMetGlyLe---AsnProAsnAspTyrIleProGlu 141
Db 3142 AAAGTGAAGAGGAATCCAGCTGCTGGCCAGCGGTGGCCAGAGATGTGCTCAACCT 3201
OY 142 TyrGln-----GlyGluGlnProAsnSerGluValValAlaProProThrLeuGlu 158
Db 3202 GTCCAGAGAGCAGAGGCGAAGAACACAGAGCAGCCT----- 3240
OY 159 ProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsn 178
Db 3241 ---GAAGCGTGGGTCTGAAAGAAAGAG-----ACGATGTAGTGTGG 3279
OY 179 LysValProArgLeuLysAlaLys-----PheTyrGlnSerSerGlnSerGlyGlu 195
Db 3280 AAATAGATGCTCGAGGAGCAAAATGAGCCTTTTACACAAAGGAAGAGTGTGGGCGAG 3339
OY 196 ThrSer-----AlaIleGlySerSerHis 203
Db 3340 ACCACCCCGAAGACCTTTGAAGAAAGCTCCCAAGTCAACAGACATATGAGTCCAGAGAG 3399
OY 204 GlnLysThr-----GluProTyrAlaAsnIleLysAlaLalaLeuGluAspIle 219
Db 3400 CTGTGAACCACTGTGCAAGCCGAAACCTTAGCTGGGTAAATCA-----CAGGAGATG 3453
OY 220 ThrGlnGlnSerAlaMet-----AspLeuAsn 228
Db 3454 GTGATGGAACAGGCTATCCCGCTGACTCGGTGGAAGACCCCTACAGACAGTGAAGTAT 3513
OY 229 GlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyr 248
Db 3514 GGAAGCACGCCCC-----GTAGCCGACTTT 3557

```


PRIOR APPLICATION NUMBER: PCT/US 00/22061
 PRIOR FILING DATE: 2000-08-11
 NUMBER OF SEQ ID NOS: 4
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 1
 LENGTH: 6608
 TYPE: DNA
 ORGANISM: Homo sapiens
 PUBLICATION INFORMATION:
 AUTHORS: Naert et al.
 TITLE: Gravin, an autoantigen recognized by serum from myasthenia gravis
 JOURNAL: Curr. Biol.
 VOLUME: 7
 ISSUE: 1
 DATE: 1997-01-01
 DATABASE ACCESSION NUMBER: U01607
 DATABASE ENTRY DATE: 1997-06-26
 US-09-738-877-1

Alignment Scores:
 Pred. No.: 0.0601 Length: 6608
 Score: 124.00 Matches: 180
 Percent Similarity: 34.07% Conservative: 130
 Best Local Similarity: 19.78% Mismatches: 301
 Query Match: 2.62% Indels: 301
 DB: 10 Gaps: 45

US-09-914-168-2 (1-919) x US-09-738-877-1 (1-6608)
 QY 31 GINGHNASNP... 45
 DB 3016 CAGAAACACACCCCGGCTACTGAACT-CTGCCAGAACAGAGAGCCCGGGGC 3074
 QY 46 AAPTTHAIAIEASNGIAIAI... 65
 DB 3075 GACAGCGCTGCTAGTACGCGGAA- 3113
 QY 66 IIEGLIAIATG... 85
 DB 3114 GTGACACCTGCAGAAACCTGCGCATTTGGTTCCGAGAGAACCCGACATCTGCT 3173
 QY 86 VALVALA... 104
 DB 3174 GCTGAGACGACACAGAAATGCTGTACAGCTCTCCAGTTAACCGAC--TCCCA 3227
 QY 105 LEUGLYLEUASP... 118
 DB 3228 -----GACACCACAGAGAGGCCACTCCGTTGCAAGAGTGAAGCTGC 3272
 QY 119 -----SerLeuGluGlu 122
 DB 3273 GTACCTACATAGACAGCAGAGAGCGGACTCAAGAGTCTCCAGCAGTGGCAGAA 3332
 QY 123 LEUPHEAL... 141
 DB 3333 AAGTGAAGAGGAATCTCCAGCTGCTGGACCGGTGGGCGCAGAGATGTGCTTACAGCT 3392
 QY 142 TYRGLN-----GlyGluGlnProAsnSerGluValValProProThrLeuGlu 158
 DB 3393 GTGCACAGACAGAGCCACCAAGACCAAGAGAGAGGCT- 3431
 QY 159 PROGLULYSPROGLYLEU... 178
 DB 3432 ---GAAGCTCGGCTGTGAAGAAAGAC- 3470
 QY 179 LYSVALPROARGLEU... 195
 DB 3471 AAGAGTAGAGTCTCAGAGGCAAAACTGAGCTTTTACACAAGGAGAGCTGTGGGAG 3530
 QY 196 THRSE... 203
 DB 196 THRSE-----AlaIIEGLYSerSerHis 203

DB 3531 ACCACCCAGAAAGCTTTGAAAAAGCTCTCAAGTCACAGAGAGCATAGCTCCAGTGG 3590
 QY 204 GINLYSTH-----GluProTyraIaAsnIIEGLYValAlaLeuGluAspIle 219
 DB 3591 CTGTGAACCATTTGTCACCGCAACCTTACTGGGGTAATCA-----CAGAGATG 3644
 QY 220 THINGLUSER... 228
 DB 3645 GTGATGGAACAGGCTATCCCTGCTGCTGAGAAACCTTACAGACAGTGAACAGTAT 3704
 QY 229 GLYSERILEPROARGLEU... 248
 DB 3705 GGAACACACCC-----GTAACCGACTT 3728
 QY 249 ASPLEASPLESERILEARGAS... 268
 DB 3729 CAGGACACGACACCAACCAAGAGAGATGTGGAA-----ATCCATGAG 3776
 QY 269 LEUGLYLUPROVALTYR... 288
 DB 3777 GAGATGAGTCT-----GCATCTGCTACCCAGTCCAGGCGCAGAGAA 3818
 QY 289 ASPLYSALAPHE... 308
 DB 3819 GCAGAGGAGTCTCTGACAGAAAGAGGCGCTCCAGACACTTCCAGTTTGTGTTCCAG 3878
 QY 309 HISGLYLYSTYR... 328
 DB 3879 GAA-----GAACTAAGAACAAATCAAGATGGAAGACCTCTAGAGCAT 3923
 QY 329 PHEASPLYAR... 348
 DB 3924 -----ACAGTAAAGAGGTGCTCATG- 3950
 QY 349 VALSERLEU... 360
 DB 3951 GTATCCATCTGTCAAAAGCTGAGGGACTCAAGAGCGTGCACCATATGCTGATGAGAA 4010
 QY 361 PHEASGLI... 375
 DB 4011 ACCAAGAGCTACATTTTTCAGAGACTTGAGGGGCTCTATAGAC-----ACAGC 4061
 QY 376 LEUTH... 383
 DB 4062 ATPACAGTCACTGCGGAAAGGTCACTGAGTGGCCCTTAAGGTGAAGGACAGAA 4121
 QY 383 ----- 383
 DB 4122 GCTGATGTAAGAAAGATGCTCTTGAACCTGCAAGTCAAGTCAAGTCTCCATCC 4181
 QY 384 PROVALY... 394
 DB 4182 CCGGTGAGAGAGATGCTGTTCAAGTCAAGGAGGAGAAACAGAGAGCCAAAC 4241
 QY 395 THRVALAS... 414
 DB 4242 CATGTGAATGAAGAGAGCTTGACGCAAGACGCTTACCGTA-----TGTGAAGAG 4295
 QY 415 ILEAL... 433
 DB 4296 GTCACTAAGAGCTCTCCAGACAGTGAATGTCCATCATATGATGGGCAAGAAAGTTC 4355
 QY 433 ----- 433
 DB 4356 AGCAGTTTGAAGAGACCTCTCCCTGCTAGTCAAGAGAGGAGCATATGACCAAA 4415
 QY 434 ---GLN... 452
 DB 4416 ATTCAAGTTCAGAGCTCTGAGGATCATCTTACACAGCGGCTGCAAGAGAGGAAAG 4475
 QY 453 ALAGIN... 472
 DB 4476 GTCTTAGAGAAACTGCCAAACATTTTAGAAACAGTGAAGCTTGAAGCTGCAAGTCA 4535

```

QY 473 ILEuMeCAspILeSerProILeGluPheSerAlaSerAsnLeuILeGlnAspLysLeu 492
    |||||
Db 4536 CATTAGTT-----CTGAGAGGAATCTCTCGAAAAAATAAGAC----- 4577
QY 493 AsnLeuValAlaAlaLysAlaArgHisLeuTYrAspMetProAspAspArgValLeuAla 512
    |||||
Db 4578 -----TTGGCCGCTCAT-----CCAGGGGAGATGCTGTCGCC 4610
QY 513 ILeAsnHisAspAspGlyValAsnArgSerILeLeuGlyArgILeSerAspAlaValSer 532
    |||||
Db 4611 ACAGGGCCCGAGCTGTGAGGCAAAATCGACACAGTG-----ATAGTATCT 4655
QY 533 AlaValAlaArgAlaILeLeuProAspGluSerGluAsnGluValILeAspLeuProGlu 552
    |||||
Db 4656 GCTACTACCAAGAAAGGCTTAAGTCCGACCTGGAAGAGAG-----AAA 4700
QY 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTYrGlnSer--LysLysVal 571
    |||||
Db 4701 ACCACATCAGTGAAGTGCAGATGAGATGACAGCTTGCTGCCAGAGAGTTC 4760
QY 572 ProLeuTYrValPheValAlaSerAspLysProArgAspGlyGlnILeGlyLeuGlyTTP 591
    |||||
Db 4761 AAGGTAGCTGTGCAATTGAGGATTAGACCTGAAATGGATTGGAACCTTGAGACC 4820
QY 592 GlySer----- 593
Db 4821 AAAAGCATAAAGCTGTCCAAACATCATCCAGACAGCCGTTGACCATTTGACTGACA 4880
QY 594 ---AspThrGlyTYrArgLeuValThr-----LysPheGlnHisAsnLeuIleAsn 609
    |||||
Db 4881 GAAGAAACAGCCACCAATGTTGACCTGTGACTTACAGACACAGCTCAGCTGATAAAA 4940
QY 610 ArgAspGlyTYrGlnAlaGlyValGluLeuArgLeuSerGluAspLysLysGlyValLys 629
    |||||
Db 4941 GCTGACAGCCAGAGAGCTGAGAGAAAGCAAGAAAGAGAGAGAACCTGAGGCTCT 5000
QY 630 LeuTYrAlaThrLysProLeuSerHisProLeuAsnAspGluLeuArgAlaThrLeuGly 649
    |||||
Db 5001 GCACAGGATGAACACCAATTCCTTCAGCCAAAGAGAGCTCAGAGTCAACCCAGTGGGA 5060
QY 650 TYrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeu-----SerThr 666
    |||||
Db 5061 CAA-----GCACATCTCATATTTCCTCCAAAGACATGAGTGAACCTTCAGAA 5105
QY 667 ArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsn-GlyLysTTPAsnArgH 686
    |||||
Db 5106 AAGACCATGACTGTGAGGTAGAGGTTCTCACTGTAATGATCAGCAGCTGGAAGAGGTC 5165
QY 686 TYrIser---LeuArgTYrArgLeuAspLysLeuLysThrGlnAlaProGluTYrThr 705
    |||||
Db 5166 GTCTCCCATCTGAGAGAGAGGAGGTGAGCTGGA-ACAAAGCTCTGCTCCAGAA----- 5219
QY 705 PGIAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaG 725
    |||||
Db 5220 -----GATGATGCTCATGCTTCTGTTAGCA-- 5243
QY 725 yValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTYrArgG 745
    |||||
Db 5244 -----GAAAGATAGAGAGAGTCACTAGTGAACCGAAAGAGAT----- 5282
QY 745 nArgTYrSerLeuGlnValGlySerSerGlyLeuValSerAsp-----AlaAsnHe 762
    |||||
Db 5283 -----GAAAAAGGTGATGAT-----GTTGATGACCTGAAAAACCAAGACTGC 5323
QY 762 tAlaIleAlaArgAlaGlyLysSerGly 771
    |||||
Db 5324 AGCCCTGCTGATCTGATGATGCTCAGGA 5351

```

RESULT 40
 US-09-880-107-3439
 : Sequence 3439, Application US/09880107
 : Patent No. US20020142981A1

```

; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3439
; LENGTH: 6608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U81607
US-09-880-107-3439

Alignment Scores:
Pred. No.: 0.0601
Score: 124.00
Percent Similarity: 34.07%
Best Local Similarity: 19.78%
Query Match: 2.62%
DB: 10
Gaps: 45

US-09-914-168-2 (1-919) x US-09-880-107-3439 (1-6608)

QY 31 GlnGlnAsnAspProAlaAsnIleLeuAsnHisValProAlaHis----- 45
    |||||
Db 3016 CAGAGAAAGAACCCCGACGGTTACTGAACTTCTCCAGAAACAGAGAGCCCGGGGC 3074
QY 46 AspThrAlaIleAsnGlnAlaLysAlaGlyAsnProValLeuLeuThrProGluGln 65
    |||||
Db 3075 GACACGGCTGTTAGTAGGCGGAA-----TTGACCCCGGAGACT 3113
QY 66 ILeGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAsp 85
    |||||
Db 3114 GTGACAGCTGTCAGAAACTGACGGCATTTGGTTCGAGAGAAACGAGACATCTGCT 3173
QY 86 ValValAsnPheAspAspLys---SerProIleSerArgIleGlyLysIleSerProPro 104
    |||||
Db 3174 GCTGAAGAGAGCAAGAAATGTTGTCAGCAGTCTCCAGATTAAACGAC---TCCCA--- 3227
QY 105 LeuGlyLeuAspMetSerValIleGlnGluTYrThrProLeu-----SerLeuGlnGlu 118
    |||||
Db 3228 -----CACACACAGAGAGGACCTCCGCTCAGAGAGTGAAGTGGC 3272
QY 119 -----SerLeuGlnGlu 122
    |||||
Db 3273 GTACTGACATAGAGAGCAAGAGACGCGACTCAAGAGCTCTCCAGAGTGGCAGAA 3332
QY 123 LeuPheAlaGlnGlnSerThrGlnMetGlyIle---AsnProAsnAspTYrIleProGlu 141
    |||||
Db 3333 AAAGTGAAGAGGAATCCAGCTGCTGCGACCGGTGGCCAGAAATGTTGCTTACGCT 3392
QY 142 TYrGln-----GlyLysIleProAsnSerGlnValValProThrThrLeuGlu 158
    |||||
Db 3393 GTCCAGAGAGAGAGAGCAAGAAACAGAGAGAGCT----- 3431
QY 159 ProGluLysProGlyLeuIleLysArgLeuTYrAlaArgLeuPheAsnAspGlyValAsn 178
    |||||
Db 3432 ---GAAGCGTGGGTCTGAGAAAGAG-----ACGATCTAGTCTTG 3470
QY 179 LysValProArgLeuLysAlaLys-----PheTYrGlnSerSerGlnSerGlyGlu 195
    |||||
Db 3471 AAGATGATGCTCAGAGAGCAAAACTGACCTTTTACACAGAGGAGAGGTGCTGGGCGAG 3530

```

QY 196 ThSer-----AlaIleGlySerHis 203
 Db 3531 ACCACCCGAAAGCTTTGAAAAAGCTCTCAAGTCAAGACAGACATGAGCTCAGAG 3590
 QY 204 GlnLysThr-----GluProTyrAlaAsnIleLysAlaIleGluAspIle 219
 Db 3591 CTGTGACCTCTGTCAGGCGAAACCTTAGCTGGGTAAATCA-----CAAGAAATG 3644
 QY 220 ThrGlnIleuSerAlaMet-----AspLeuAsn 228
 Db 3645 GTGATGGAACAGGCTATCCCTCGACTCGGTGGAACCCCTACAGACAGTGAAGCTGAT 3704
 QY 229 GlySerIleProArgLeuArgGlnThrAlaLeuValAlaIleArgAlaValGlyTyr 248
 Db 3705 GGAAGCACCCTC-----GTAGCCGACTTT 3728
 QY 249 AspIleAspIleuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAsp 268
 Db 3729 GACGACCCAGGACACACCCAGAAAGAGAGATGTTGTGAA-----ATCCATGAG 3776
 QY 269 LeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAsp 288
 Db 3777 GAGATGAGCTC-----GCATCTGTACCCAGTACAGGGGACACAGAA 3818
 QY 289 AspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis 308
 Db 3819 GCAGAGCCAGTCTCTGACAGAAAGAGAGGCTCCAGCACCCTTCCAGCTTTGTGTTCCAG 3878
 QY 309 HisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnIleSerAlaGlnHisGlyTyr 328
 Db 3879 GAA-----GAACTTAAAGACATCAACATGAGACACCTAGAGCAT----- 3923
 QY 329 PheAspGlyArgTyrPheAspArgSerValAspValIleLeuProAspAsnThrAlaAsp 348
 Db 3924 -----ACAGATAAAGAGCTGCTCAGTG-----GAAACT 3950
 QY 349 ValSerLeuIleTyrAspThr-----GlyThrGln-----TyrArg 360
 Db 3951 GTATTCATCTGTCAAAGACTGAGGGACTCAAGAGCTGACAGATGCTGATGAGACA 4010
 QY 361 PheAspGluValAlaPhePhe-----ThrIleAspProLysThrAsnGln 375
 Db 4011 ACCAAGAGCTACCATTTTTCGAAGACTTGAAGGGGTCTATAGC-----ACAGGC 4061
 QY 376 LeuThrThrAspProAspLysLeu----- 383
 Db 4062 ATAAACAGTACGCGGAAAGTCACTGAAGTTGCCCTTAAGGTGAAGGACAGAAAGAA 4121
 QY 383 ----- 383
 Db 4122 GCCTGAATGTAAAAAGATGATGCTTTGAACGTCAGAGTCACGCTAAGTCTCCTCATCC 4181
 QY 384 ProValLysArgGluLeuLeuGlnLeu-----Leu 394
 Db 4182 CCCGTGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4241
 QY 395 ThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeu 414
 Db 4242 CATGTGAATGAAGAGAGCTTGAGACAGAAAGCTGTTACCTA-----TCTGAAGAG 4295
 QY 415 IleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGlu 433
 Db 4296 GTCAAGTACAGCTCTCCAGACAGATGATGCTGATGATGATGATGATGATGATGATGATGAT 4355
 QY 433 ----- 433
 Db 4356 AGCAGTTTGAAGGAGACCTCTCTCTGCTAGTCAAGACAGAGCAGTATGACACAAA 4415
 QY 434 ---GlnIleGlnAsnAspGlnValSerPheGluGlnIleSerSerSerArgTyrGluPro 452
 Db 4416 ATTCAAGTTTCAAGAGCTCTGAGGATCATTCATCTTAACAGCGGCTGACAGAGGAAAG 4475
 QY 453 AlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAspGly 472

Db 4476 GTCTTAGAGAAACTGGCAACATTTTAGAAAACAGGTGAACAGCTTGAGCCTGCAGTGCA 4535
 QY 473 IleuMetAspIleSerProIleGluPheSerIleAsnLeuIleGlnAspLysLeu 492
 Db 4536 CATTTAGTT-----CTGGAAAGAGAAATCCTGTAATAAATAAAGAAC----- 4577
 QY 493 AsnLeuValAlaIleLysAlaArgHisLeuTyrAspMetProAspAspArgValLeuAla 512
 Db 4578 -----TTTCCGCTCAT-----CCAGGGAGAGTCTGTCCC 4610
 QY 513 IleAsnHisAspAspGlyValAsnArgSerIleLeuGlnArgIleSerAspAlaValSer 532
 Db 4611 ACAAGCCCGCCAGCTGTGAGCAAAATGACACACAGT-----ATAGTATCT 4655
 QY 533 AlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
 Db 4656 GCTTACTACCAAGAAAGCTTAAGTTCGACCTGGAAGAGAG-----AAA 4700
 QY 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSer-----LysLysVal 571
 Db 4701 ACCACATCACTGAAGTGAAGTCAAGATGATGATGATGATGATGATGATGATGATGATGAT 4760
 QY 572 ProLeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyr 591
 Db 4761 AAGTGAGTGTAGCAATGTGAGGATTTAGACCTGTAATAATGGATTTTGAACCTTGACACC 4820
 QY 592 GlySer----- 593
 Db 4821 AAAAGCTAAACTGTTGTCAAACATCATCCAGACACCCCTTGACCACTTGTAGCTACA 4880
 QY 594 ---AspThrGlyThrArgLeuValThr-----LysPheGlnHisAsnLeuIleAsn 609
 Db 4881 GAGAAACACCCACCGAAATGTTGAGCTGTGATGATGATGATGATGATGATGATGATGATGAT 4940
 QY 610 ArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysGlyValLys 629
 Db 4941 GCTGACGCGCAGACGCTGAGAGAAACGAGAAAGAGAGAGAACTCAGGCTCT 5000
 QY 630 LeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuValArgAlaThrLeuGly 649
 Db 5001 GCACAGAGTAAACACCAATTAATCTTCAACAAAGAGGAGCAAGTCAACCCAGTGGGA 5060
 QY 650 TyrGlnGlnGluValPheGlnHisSerThrAsnGlyPheAspLeu-----SerThr 666
 Db 5061 CA-----GCACATTCGTGATATTTCCAAAGACATGATGAAGCTCAGAA 5105
 QY 667 ArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsn-GlyGlyTyrPasnArgTh 686
 Db 5106 AAGACCATGACTGTTGAGTGAAGAGTTCACCTGTAATGATGATGATGATGATGATGATGATGAT 5165
 QY 686 rTyrSer-----LeuArgTyrArgLeuAspLysLysThrGlnAlaPheProGluThrTr 705
 Db 5166 GTCTCCCATCTGAAGAGAGGAGGTGAGCTGGA-ACAAAGCTCTGTGCGAA----- 5219
 QY 705 pGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaG 725
 Db 5220 -----GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5243
 QY 725 yValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgG 745
 Db 5244 -----GAAAGATAGAGAAAGTCACTAGTTGAACCCAGAAAGAT----- 5282
 QY 745 nArgTyrSerLeuGluValGlySerSerGlyLeuValSerAsp-----AlaAsnMe 762
 Db 5283 -----GAAAGAGTATGAT-----GTTGATGATGATGATGATGATGATGATGATGATGAT 5323
 QY 762 tAlaIleAlaArgAlaGlyLysSerGly 771
 Db 5324 AGCCCTGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5351

Search completed: May 6, 2003, 02:59:34

Tue May 6 09:37:18 2003

us-09-914-168-2.rmpb

Page 65

Job time : 556 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 5, 2003, 23:56:45 ; Search time 2019 Seconds

(without alignments)
7371.794 Million cell updates/sec

Title: US-09-914-168-2
Perfect score: 4727

Sequence: 1 MSKPVLPANRSMFVALAAY.....TGVEKGNPKLHFHFIGTPE 919

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1654066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgmt2_1/USPRO.spool/US09914168/runat_28042003.151615.6799/app_query.fasta_1.1095
-DB=EST -QFMT=fastap -SUFFIX=trst -MINMATCH=0.1 -LOOPEXT=0 -LIST=45
-UNITS=b1ts -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -MODE=LOCAL
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=40 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09914168 @CGN.1.1.1477 @runat.28042003.151615.6799 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEDUQUEY -NEG.SCORES=0 -WAIT -LONLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

EST :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlmu:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	153	3.2	792	14	B0506223	B0506223 EST613638
C 2	143	3.0	649	17	BH396320	BH396320 AG-ND-161
C 3	129.5	2.7	620	12	BF942509	BF942509 1D2 CDNA
C 4	128.5	2.7	423	9	A1363592	A1363592 SMOVL2CAS
C 5	128.5	2.7	503	10	BE459389	BE459389 EST414681
C 6	127.5	2.7	3404	11	AK004785	AK004785 Mus muscu
C 7	124	2.6	2049	11	AT161466	AT161466 Homo sapi
C 8	121.5	2.6	494	10	AW191135	AW191135 68702D06
C 9	118.5	2.5	3897	17	BH770994	BH770994 LMGtag71
C 10	117.5	2.5	2582	11	BC011226	BC011226 Mus muscu
C 11	117.5	2.5	2607	11	BC032205	BC032205 Mus muscu
C 12	117.5	2.5	2837	11	BC032614	BC032614 Homo sapi
C 13	116	2.5	607	10	AV557254	AV557254 AV557254
C 14	113.5	2.4	503	11	AY108685	AY108685 Zea mays
C 15	112	2.4	544	9	A1519730	A1519730 LD39445.5
C 16	111	2.3	3144	11	BC025329	BC025329 Homo sapi
C 17	110.5	2.3	678	12	BF491509	BF491509 AT28265.5
C 18	110	2.3	1673	11	BC013085	BC013085 Homo sapi
C 19	109	2.3	1501	11	AY104509	AY104509 Zea mays
C 20	107	2.3	2083	11	BC022490	BC022490 Homo sapi
C 21	106	2.2	588	10	AW053211	AW053211 614073B08
C 22	105	2.2	758	12	BE786630	BE786630 601475986
C 23	105	2.2	1557	107	BH770729	BH770729 LMGtag47
C 24	104.5	2.2	1021	17	CNS071K5	AL425271 clone BAO
C 25	102.5	2.2	1364	10	AW342205	AW342205 GtHST91
C 26	102.5	2.2	1867	11	BC019335	BC019335 Homo sapi
C 27	102.5	2.2	1876	11	BC018883	BC018883 Homo sapi
C 28	102	2.2	1014	13	B1086482	B1086482 602849752
C 29	102	2.2	2039	11	AK013402	AK013402 Mus muscu
C 30	101.5	2.1	513	17	AZ927049	AZ927049 476.d1s16
C 31	101	2.1	2340	11	BC028159	BC028159 Homo sapi
C 32	101	2.1	2336	11	AK011769	AK011769 Mus muscu
C 33	100.5	2.1	912	9	AL546961	AL546961 AL546961
C 34	100	2.1	800	12	BE256095	BE256095 HVSMEF000
C 35	100	2.1	1041	9	AL523772	AL523772 AL523772
C 36	100	2.1	3499	11	AK019509	AK019509 Mus muscu
C 37	99.5	2.1	969	17	CNS013XL	AL103443 Drosophi1
C 38	99.5	2.1	1514	11	AY108441	AY108441 Zea mays
C 39	99.5	2.1	1562	11	AK002434	AK002434 Mus muscu
C 40	99	2.1	1022	14	W29713	W29713 mc05a09.f1
C 41	99	2.1	1970	11	AK010097	AK010097 Mus muscu
C 42	99	2.1	2345	11	AK014398	AK014398 Mus muscu
C 43	99	2.1	4379	11	BC008524	BC008524 Mus muscu
C 44	98.5	2.1	884	17	BH137735	BH137735 ENTNI81TF
C 45	98.5	2.1	1409	11	AY109022	AY109022 Zea mays

ALIGNMENTS

RESULT 1
B0506223/c
LOCUS B0506223
DEFINITION B0506223 792 bp mRNA linear EST 22-JUL-2002
analysis mixed potato tissues Solanum tuberosum CDNA clones for microarray
3' end, mRNA sequence.
ACCESSION B0506223.2 GI:21922105
VERSION B0506223.2
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; eusteids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 792)

us-09-914-168-2.rst

AUTHORS

Other_ESTS: EST613637

SOURCE	Location/Qualifiers
1	700

724 AAGLYVALAVALAHISISTHRYVALA-

FEATURES

SOURCE

2.

[illegible]

TITLE	JOURNAL	COMMENT
Direct Submission		Carninci,P., Baldarelli,R.,Bono,H., Brownstein,M., Gult,C., Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hoti,F., Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K., Numasaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Ouackenbush,J., Salto,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Saeki,D., Schiraldi,L., Shibata,K., Shibuta,Y., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Submitted (10-JUL-2000)		Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Shuto-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-resgsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
		Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
		cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGACGAGGAGCGCCGCACACGATGGATTTTCTTTTTTNN 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence[5' GAGAGCAGAAGCATCAAGACTCAATTAAATTAATAACC(CCCC(C)C) 3'. cDNA was cleaved with XhoI and SfiI. Cloning sites, 5' end: Ssti; 3' end: XhoI. Host: SOLR.
FEATURES	SOURCE	
		Location/Qualifiers
		1..3404
		/organism="Mus musculus"
		/strain="C57BL/6J"
		/db_xref="FANTOM_DB:1200015F06"
		/db_xref="MGM:MGI:1904368"
		/db_xref="taxon:10090"
		/clone="1200015F06"
		/sex="male"
		/tissue-type="lung"
		/clone_lib="RIKEN full-length enriched mouse cDNA library"
		/dev_stage="adult"
		1..3404
		/gene="Myptl"
		1..3015
		/gene="Myptl"
		/note="data source:MGD, source key:MGI:1309528, evidence:ISS
		myosin phosphatase, target subunit 1 putative"
		/ccoon_start=1
		/protein_id="BAB23563.1"
		/db_xref="GI:12836228"
		/db_xref="MGD:MGI:1309528"
		/translation="MKMDAKOKRNFDLKNIGSEFTLEPPVVKROKTKVPDFGAVFLAACSSGDTDEVLLKLRGADINRYANDGLALHQAQCIDDNVDVYKFLVENGINPDPNEWDIPARAAECGVLDIAETLIGQAHVGVAWSBEDPTLDIAEEAMEELLONEVNRFGCDIEAPKAAREERVMI,DAERWLNSGHLSDVSHAKSGGTALHAHVAKGYTEVLKLILIOAQDVNIKYDDWPTPLAHHAMKEEACRLIVDNCIDMETVAKVCGTAFADVDEDILGYLEELQKOTILLHSKRKKSPLESTAMENNOPOKAKKETLIIPEKNAIRITLESLEHKRADEEBEEKDDESSCSSEDEEDDESAPETDKTPMAVSYNATHSTQOAPAAVYAPLTSSMOQTPIPSVKKFPISITISKPEEKERDESPASWRDLRTKGSTGALA EISAKEAKOEKTDAGVMRASSPRLSSSLDNKREKDMKTRLAYVTPTIPRLASTSIDIEKNRESSLSRTSYSTRKMEDLKNKSINGSGTYHRCSFGRODDLSCSVPTSTPTVTSAGLORSLSNSTAKPFGSSAGTOSNTRLMABDSFEKDK SAPATLVPAVYVNAAPSTTLITLTGTAGTVSEVERRRSYLTPVDESSESORAE RSROARSRSNGCVTLTDLQEAKTIGRSSTTRQDENBEKKEKEKODKOE KKEEAREREDEYKORYSRITYDETTRYRPVSTSSSAASSSLSTLSTLYASSQLNR
gene	CDS	

QY	651	GINGINGLIVALPheGLyHisSerThrsn-----GlyPheAspLeuSerThrArg	667
Db	1534	GAAGAGAAAGACAGACAGAGCTTCAAGTTTGCGAACACAAGTAGTTCTTACCAAGAAGA	1593
QY	668	THrLeuGLUHSISLeuSerArg-----SerIleIeGLnAsnGLyTYRpaSnaRG	685
Db	1594	AATGGAGAAAGATGACTTAAAAGAATAGTTCGCATCATGAAGAGATCCACTTACCATAAGA	1653
QY	686	THrTYRSerLeuArGTyrArgLeuAspLysLeuLSThGlnAlaProProGLuthr	704
Db	1654	ACTTCCTCCTTGTGTAGAACAGATGATTGATTAGTTGATGAGTTCCAAAGCAC	1710
RESULT 7			
LOCUS	AF161466	2049 bp	mRNA linear HTC 22-MAY-2001
DEFINITION	Homo sapiens HSPc117 mRNA, complete cds.		
VERSION	AF161466		
KEYWORDS	HTC.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 2049)		
JOURNAL	Zhang, Q.H., Ye,M., Wu,X.Y., Ren,S.X., Zhao,M., Zhao,C.J., Fu,G.,		
PUBLISHED	Shen,Y., Fan,H.Q., Lu,G., Zhong,M., Xu,X.R., Han,Z.G., Zhang,J.W.,		
REFERENCE	Tao,J., Huang,Q.H., Zhou,J., Hu,G.X., Gu,J., Chen,S.J. and Chen,Z.		
AUTHORS	Cloning and functional analysis of cDNAs with open reading frames		
TITLE	for 300 previously undefined genes expressed in CD34+ hematopoietic		
JOURNAL	stem/progenitor cells		
PUBLISHED	Genome Res. 10 (10), 1546-1560 (2000)		
REFERENCE	2 (bases 1 to 2049)		
AUTHORS	Ye,M., Zhang,Q.H., Zhou,J., Shen,Y., Wu,X.Y., Guan,Z.Q., Wang,L.,		
TITLE	Fan,H.Y., Mao,Y.F., Dai,M., Huang,Q.H., Chen,S.J. and Chen,Z.		
JOURNAL	Human full length cDNA cloned from cd34+ stem cells		
PUBLISHED	3 (bases 1 to 2049)		
REFERENCE	Ye,M., Zhang,Q.H., Zhou,J., Shen,Y., Wu,X.Y., Guan,Z.Q., Wang,L.,		
AUTHORS	Fan,H.Y., Mao,Y.F., Dai,M., Huang,Q.H., Chen,S.J. and Chen,Z.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-JUN-1999) Shanghai Institute of Hematology, Shanghai		
PUBLISHED	Second Medical University, Rui-jin Hospital, 197 Rui-jin Road II,		
REFERENCE	Shanghai 200025, P. R. China		
AUTHORS	Location/Qualifiers		
TITLE	1.. 2049		
JOURNAL	/organism="Homo sapiens"		
PUBLISHED	/db_xref="taxon:9606"		
REFERENCE	/clone="CBCHC01"		
AUTHORS	/cell_type="cd34+ stem cells"		
TITLE	/tissue_type="blood"		
JOURNAL	91.. 1608		
PUBLISHED	/codon_start=1		
REFERENCE	/product="HSPc117"		
AUTHORS	/protein_id="AAF29081.1"		
TITLE	/db_xref="GI:6841456"		
JOURNAL	/translation="MSRSYNDDELFLFKINKRCWRKIKKGFVPMNQVEGVFYNDALEKK		
PUBLISHED	LMEFEALNKGSGVGGLPALMKOIGNVALPGIVHSIGLDVHGSGFGAIGMAALE		
REFERENCE	DMDPEAVASPGGVDFNGCVRLIRLTDSDDPYKEOLAAMPDHIPVGSCKSV		
AUTHORS	IPMAKDLFEALEMGVMSLREGYAMEDKCHCEHYGMLOADPNKYSAKRAKKGLPQ		
TITLE	LGTLGACNHRYAEIOVDIEFNFEPAKKMGIDHKQVCWPHISGRGLGHQVATDALVA		
JOURNAL	MEKAKMRDKIIIVNRQLALSSNASPEVODYLKGAAGNVAVNRRSMTELTRQAFAR		
PUBLISHED	VFTNPDDLDTLVITYDVSHNIKAQVOHVNDKERETLTVHRKSTRAPRPHPHLIAVDYD		
REFERENCE	OVTPCPVLIGTMTGTSYVLTGTGQTGMTGTCTTCHGAGRALSRASKSRINDFDVDL		
AUTHORS	KLADMGILRIVASPLFIWEAPESKNTVDVNTCHDAGISKALKIKRLPIAVING"		
BASE COUNT	553 a 435 c 562 g 499 t		
ORIGIN			

	Percent Similarity:	34.46%	Conservative:	71
	Best Local Similarity:	21.16%	Mismatches:	204
	Query Match:	2.62%	Indels:	146
DB:		11	Gaps:	25
US-09-914-168-2 (1-919) x AFL161466 (1-2049)				
QY 161	LysPrgIleuLiLeIysAtrgLeuTYrAla-----ArgIeuPheAsnsp-----	175		
Db 58	GAAGCCGGAGATTGTTTGGCTACAGTTTGGCCACCATGATGGCAGTATATATATGAGCTG	117		
QY 176	-----GlyValAsnLys--ValPrgIeuLysAlaLysPheTYrgInser	190		
Db 118	CAGTTCCTTGAGAGATCAATATAAACTGCTGGAGGATCAAGAAAGGCTTCGCCAAC	177		
QY 191	SerGInserGlyIuThrSerAlaIleGlySerSerHisGInLysThrGluProTYrAla	210		
Db 178	ATGCAG-----CTTGAAGTGTTC	198		
QY 211	AsnIleLysAlaAlaLeuGluAspIleThGInGInserAlaMetCsp-----	226		
Db 199	TATGTGATGATGCTCTCGAGAAATTATGTTTGAGGAATTAAGATGCTTCGAGCT	258		
QY 227	-----LeuAsnGlySerIleProArGleuArGInThrAlaLeuValAlaAlaArGala	244		
Db 259	GGTGGTGTGGTGGCTTCCTCCAGCCAGTAACAGATGGCAATGTGGCAGCCGTGCT	318		
QY 245	ValGlyTYrTYrAspIleAspLeuSerIleIleArGInSerIleGlyGluValAspAl	264		
Db 319	-----GGAATGTTATCATCATATATGGCTTCCTGATGTC	354		
QY 265	-----IleIleHisAspLeuGluProAl	273		
Db 355	CATTGAGATATGGGTTTGGTATTGGAGACATGGACACTTGTATATGATGACCTGAA	414		
QY 274	TYrIleAspTYrArGAlaValaGluValArGlyGluGlyAlaSp-----Asp	289		
Db 415	GCAGTA-----GTAATCCCGAGGGTCTCGGTTTGACATCAACTGGTGTGTC	462		
QY 290	LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHis	309		
Db 463	CGCTTCCTCAACACCAATTATGATGAA-----ACTGATGTC-----	498		
QY 310	GlyIysTYrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTYrPhe	329		
Db 499	-----CAGCCTGTGAAGACCACTTGCACCACTATGTTGATGCAC	540		
QY 330	AspGlyArGTYrPleuAspArgSerValAspAl-----IleLeuPro	343		
Db 541	-----ATTCTCTTGGCGTGGGCTCAAAAGCTGATCAACCA	576		
QY 344	AspAsnThrAlaAspValSerLeuIleTYrAspThrGlyThrGInTYrArGPhaSpGlu	363		
Db 577	ATGATGGCAAAAGACTTGGAGAGCCTTGGAGATGGGGTGGACGTGCTTAAAGAA	636		
QY 364	ValValPhePheThrIleAspProLys-----ThrAsnGInLeuThrThr	378		
Db 637	GGGTATGCTGGCGTGAAGACAGACACTGCGAGAGATACGAGAGATGCTGCAGCT	696		
QY 379	AspProAspLysLeuPro-----ValLysArGluLeuLeuGInGInLeuThrVal	396		
Db 697	GAACCCCAATAAGTTTCTGCAGAGGCGCAAGAAAGAGGCTTCTCTCAGTTGGGACCTTG	756		
QY 397	AsnMetGlyGluAlaTYr--AsnLeuGInAlaValArGAlaLeuSerAsnAspLeuIle	415		
Db 757	GGAGCAGCAGCAACATTATGAGAAATCCAGAGTGGATGATGATTTTCATGAGTATGCT	816		
QY 416	AlaThrArGTYrPheAsnMetValAsnThGInLeuIleValPheProGluArGluGInIle	435		
Db 817	GCTAATAAAA-----ATGGCATGCAC-----	837		
QY 436	GlnAsnAspGluValSerPheGluGInLysSerSerArGInThrGluProAlaGInVal	455		

Genetique Microbienne

INRA

CRI INRA, Domaine de Villvert, 78352 Jouy en Josas cedex, France

Tel: 33 1 34 65 25 16

Fax: 33 1 34 65 25 21

Email: sotokine@jouy.inra.fr

best homologue in strain IL1403 is dnaE (92%)

Class: shotgun

High quality sequence start: 30

High quality sequence stop: 3867.

FEATURES

Source

Location/Qualifiers

1. 3897

/organism="Lactococcus lactis subsp. cremoris"

/db.xref="taxon:1359"

/clone.lib="MGI363 Random Sequence Tag Library"

/note="Vector: pSGM2; Site 1: SmaI; Library of chromosomal fragments of L.lactis strain MGI363 was prepared by partial AluI digestion or by sonication."

BASE COUNT 1211 a 628 c 848 g 1210 t

ORIGIN

Alignment Scores:

Pred. No.:	0.0834	Length:	3897
Score:	118.50	Matches:	122
Percent Similarity:	34.35%	Conservative:	114
Best Local Similarity:	17.76%	Mismatches:	236
Query Match:	2.51%	Indels:	215
DB:	17	Gaps:	32

US-09-914-168-2 (1-919) x BH770994 (1-3897)

QY 199 IleglSerSerHlglnLysThrGluProTyrAlaAsnIleLysAlaLaleuGluasp 218
 ||||| : : : : : |||||
 Db 1680 ATTGGTTTACAGAGAAGAGAGAGAAATGTTT-----GCACCTTAATACA 1727
 QY 219 lIethrGlnJuserAlaMetaspLeuasnGlySerIleProArgLeuArgGlnThrAla 238
 ||||| : : : : : |||||
 Db 1728 AAAACAGATACAGTCTTTTATGAT-----ACTGTTGTTAACTA---GACGATTAT 1775
 QY 239 LeuValAlaAlaArgAlaValAlaGlyTyrTyrAspIleaspLeuSerIleleArgasn--- 257
 ||||| : : : : : |||||
 Db 1776 TTGGAGAGGGCTTACGTTTACGCTATAAACGTCGGAATTTGATGGGAAACCTC 1835
 QY 258 -----SerIleGlyLalValaspAllelle 266
 ||||| : : : : : |||||
 Db 1836 CATGACGCTTTGGTTGTCAGAAAGCTCAATGCTTAATTACACACACTTATTAGT 1895
 QY 267 HisaspLeuGlyGluProValTyrIleaspTyrArgAlaValGluValArgGlyGluGly 286
 ||||| : : : : : |||||
 Db 1896 ATTGCATTAATAAT-----TTTGAATGCGGGGTTGCCAGTT----- 1931
 QY 287 AlaaspAspLysAlaPheThrThrValAlaaspGluValProLeuLeuIleGlyaspVal 306
 ||||| : : : : : |||||
 Db 1932 -----TCCTTTTCTTTTATTCCTTAATAATACA----- 1958
 QY 307 PheHISglLysTyrGluThrLysAsnLeuIleGluAsnAlaSerAlaGluHis 326
 ||||| : : : : : |||||
 Db 1959 -----GAAAGTTATATAAAACTTCTT---CGGATTTCACACTTTCAT 1997
 QY 327 GlyTyrPheaspGlyArgTyrLeuasp-----ArgSerValaspValle 341
 ||||| : : : : : |||||
 Db 1998 AATTATGATCAGCGCATTTTTCGATATTCGAAATCATTTTGTCCAGATGCCATGGGT 2057
 QY 342 LeuProaspAsnThrAlaaspValSerIleuIleTyrAspThrGlyThrGlnTyrArgPhe 361
 ||||| : : : : : |||||
 Db 2058 ATTCCAGAACTTATGGAGCTTTAGTCAGATCAGAGCTTCACGA-----TTT 2108
 QY 362 AspGluValValPhePheThrIleaspProLysThrAsnGlnLeuThrThrAspProasp 381
 ||||| : : : : : |||||
 Db 2109 GCCCAGCAAGGCTATATTGGATTACACGACGACATATAA----- 2150
 QY 382 LysLeuProValLysArgGluLeuLeuGluLeuLeuThrValAsnMetGlyGluAla 401

Db 2151 -----GGAACCTAAA 2159
 QY 402 TyrAsnLeuGlnAlaValArgAlaLeuSerAsnaspLeuIleAlaThrArgTyrPheasn 421
 ||||| : : : : : |||||
 Db 2160 TTTAATTTCACAAACCTT-----CCTTTCCAGAGTTCGTTATCTGAT 2204
 QY 422 MetValAsnThrGluIleValPheProGluArgGluGlnIleGluAsnaspGlnValSer 441
 ||||| : : : : : |||||
 Db 2205 TTTGCTGATATGAAATTTTA----- 2225
 QY 442 PheGluGlnSerSerSerArgThrGluProAlaGlnValaspLuserThrLeuGlu 461
 ||||| : : : : : |||||
 Db 2225 ----- 2225
 QY 462 ProValIleGluThrValGluLeuThrAspGlyIleLeuMetaspIleSerProIleGlu 481
 ||||| : : : : : |||||
 Db 2236 GCTATTTTGCAGCGCGTGCCT-----GATGGGCTTTCCTTTGACGAAAGT---TTGACG 2276
 QY 482 PheSerAlaSerAsnLeuIleGln----- 489
 ||||| : : : : : |||||
 Db 2277 GTCAGTTCAATGAATTCCTCCAAAGCCACAGACTTATGAATCTTATTTCCAGAAATAT 2336
 QY 490 -----AspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAsp 504
 ||||| : : : : : |||||
 Db 2337 TTTCGCAAGCACGTGCAAAATCTGCAGTATTAACAGCAAAATTCGTTATCAATTT--- 2393
 QY 505 MetProaspAspArgValLeuAlaIleAsnHisaspaspGlyValAsnArgSerIleLeu 524
 ||||| : : : : : |||||
 Db 2394 -----GATGAAATAATTCGATCCGAGATTTGATTAACAAAGGAAACGACCTTTGAA 2444
 QY 525 GlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProaspLuserGlu 544
 ||||| : : : : : |||||
 Db 2445 CATCTCCAGAGAAGACGATTTTAGACTCTGCTCCGACTTGAAGAAAAAATCACTG 2504
 QY 545 AsnGluValIleaspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlasp 564
 ||||| : : : : : |||||
 Db 2505 ACAGAAATTTGCTGATGAAGTCTGCAAGTCTTTAGCGTGAAGATCAACGACGATGAT 2564
 QY 565 -----ValTyrGln-----SerLysLysValProLeu----- 573
 ||||| : : : : : |||||
 Db 2565 TTTCAGTATATACAGACGCTTTAGAAAAAAGAGTTTCAGTCATTCACCAAAATGGGATTT 2624
 QY 574 -----TyrValPheValAlaSerasp-----LysProArgaspGlyGln 586
 ||||| : : : : : |||||
 Db 2625 GATGATATTTTTCATGTCGCGGATTTTTCGCCCTATGCCGTCGACAAACACATTTAT 2684
 QY 587 IleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeu----- 599
 ||||| : : : : : |||||
 Db 2685 TCGGGAATGGGCGTGGCTGCGCTGCTGGCTTGGTTCCTATGTTTGGGATTCAC 2744
 QY 600 ---ValThrLysPheGluHisAsnLeu-----IleAsnArgaspGlyTyr 613
 ||||| : : : : : |||||
 Db 2745 CAGGTTGACCCCTGCAAAATATCTTTCTTTGAACGTTTTTTTGAACCCAGAAAGAGTNG 2804
 QY 614 ---GlnAlaGlyAlaGluLeuArgLeuSerGluAspLysGlyValLysLeuTyrAla 632
 ||||| : : : : : |||||
 Db 2805 GCCATGCCCGATATTGATATGATATGATGATGATGATGATGATGATGATGATGATGAT 2864
 QY 633 ThrLys-----ProLeuSerHisProLeuAsnaspGlnLeuArgAlaThrLeuGly 649
 ||||| : : : : : |||||
 Db 2865 ATGAAATAATCGTTATGTTACAGACCATGTAGACAAATGTACATTTTCAACTTTAGCA 2924
 QY 650 TyrGlnGln-----GluValPheGlyHisSerThrAsnGlyPheasp 663
 ||||| : : : : : |||||
 Db 2925 AAAACAGAGCGCTCAGAGATGTCGTAAGCCTTTGGGATGACCGAAGCTGAATTAATCT 2984
 QY 664 LeuSerThrArgThrLeu-----GluHisGlu----- 672
 ||||| : : : : : |||||
 Db 2985 AATTGACTCGGATGCTTTCGCGCTTTTGGAACTCTGCGCTGATGATGATGATGATGAT 3044
 QY 673 -----IleSerArgSerIleIleGluAsnGlyGlyTyrPasnArgThrTyrSerLeuArg 690
 ||||| : : : : : |||||


```

Oy 317 AsnLeuLeuGlu-----AsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyr 333
    ::::: ||| :::: |||
Db 907 GATCTTGTTGGGGAAGTACATGACACATTGAAGAAAGATGACAGGGGAAACCT 966
Oy 334 LeuAspArgSerValAspValIleLeuPro-----AspAsnThrAlaAspValSer 350
    ||| :::: |||
Db 967 GATCAAGATCAGAGAGGTGAAGTTTCCCTCGAAGTACAGCAGCGCCGAG----- 1020
Oy 351 LeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPheIleThrLeasp 370
    ::::: ||| ::::: |||
Db 1021 -----GAA 1023
Oy 371 ProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeu 390
    ||::: ||| ::::: |||
Db 1024 CCCGATCTTCTGATGCTGCTAGT-----GACTCTGCTTGGCCATATAAATAAGGA 1077
Oy 391 GlnGlnLeuLeuThrValAlaSerMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeu 410
    ||::: ||| ::::: |||
Db 1078 GAAAAG-----AATGAAAGGCGCAGATGTCATTGTGGA-----TTG 1113
Oy 411 SerAsnAspLeuIleLeuThrArgTyrPheAsnMetValAsnThrGluIleValPhePro 430
    ||| :::: |||
Db 1114 AAGATACTCTG-----GATATTGTTGATTAACCTGAGAACTGACCT 1155
Oy 431 GluArg-----GluGlnIleGlnAsnAspGlnValSer 441
    ::::: |||
Db 1156 CAGAGAAATGAGTGGAAACACTGGGCTATGGTGGAGCAGACGAGCTCATATGCTCAG 1215
Oy 442 PheGlnGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGlu 461
    ::::: |||
Db 1216 TTACAGAGCAGCAGATTCATAGTCAGATTGGAGGAAGTGCATTCCTGCTTTGAA 1275
Oy 462 ProValIleGlnThrValGlnLeuThrAspGlyIleLeuMetAsp----- 476
    ||::: ||| ::::: |||
Db 1276 CCGGAAGCTAGCAGCTTAAATAATCTATTGATGCTCTTACCAGAAATTCAGAAACA 1335
Oy 477 -----IleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsn 493
    ||::: ||| ::::: |||
Db 1336 TTGAATATTACTCAAGATTATAATGGAACAGCCATCTTCAGATAGTAAGTGGT 1395
Oy 494 LeuValAla-----AlaLysAlaArgHisLeuTyrAspMetProAspAsp----- 508
    ||| :::: |||
Db 1396 TTAGAGCCCAAAACATAAAACCCCAACACATAAATCTGTAATCTATTCTAAACAAGC 1455
Oy 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAsnArg 521
    ::::: |||
Db 1456 ATGACCACAGACACACCCAGAAAACCTGGCGCAGCAAGCATGAAGTACTATACAAA 1515
Oy 522 SerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAsp 541
    ::::: |||
Db 1516 ACT-----AAATCTAATGTCAGAGCTGTGACAGCA-----AAT 1548
Oy 542 GluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThr 561
    ||| :::: |||
Db 1549 TCTGATGACCCGACACCCAGCCCTGATTTCTCAGAGGCCAGTCAAACTGAGAAAAACAA 1608
Oy 562 ProAlaAspValTyrGlnSerLysValProLeuTyrValPheValAlaSerAspLys 581
    ||| :::: |||
Db 1609 GGTGATACGTTTGGAAAGCTCAG----- 1632
Oy 582 ProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuValThr 601
    ::::: |||
Db 1633 -----AGTTGTAATCTCGAGCTTAATCTGTGAAA 1662
Oy 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuLeu 621
    ||| :::: |||
Db 1663 ACCCAAGCTCATTTCTGTATTGAAA-----ACAATG 1692
Oy 622 SerGluAspLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
    ::::: |||
Db 1693 CCCAGAGATCAAAACACATGCGAGATT-----TCCAAACCTTAATCTACTCTCATAGT 1746
Oy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661

```

```

Db 1747 GATACGCTTCAT-----GATCAGCT----- 1767
Oy 662 PheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGly 681
    ::::: |||
Db 1767 ----- 1767
Oy 682 GlyTyrAsn-ArgThrTyrSerLeuArgThrArgLeuAspLysLeuLysThrGlnAlaBr 701
    ||::: |||
Db 1768 GGCTTCTTAAGAGGCTCCCTCCATCTGTACAACTGGAC----- 1807
Oy 701 oProGluThrThrGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAl 721
    ||::: |||
Db 1808 -----ACTTGG-----TGATTCAGCTCAAAA-CAGACCCAGAAACCTCAGACAGGCG 1856
Oy 721 aLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMet-- 740
    ||::: |||
Db 1857 AGCTGTGGGGAAGGCGAGTACGCCAT-----GTGAAGGATGAGCATGACCATCAGTCCAG 1910
Oy 740 ----- 740
Db 1911 TGAGCATCTTAAGAGAGCATTAACCTGAAGCCAGAAAACCTGATAGCACTGCAGCC 1970
Oy 741 -ArgGlyTyrArgGlnArgTyrSerLeuGlnValGlySerSerGlyLeuValSerAspAl 760
    ||| :::: |||
Db 1971 TCGCCAAAGAGAGAGCAGCCGAAGTTTCTCTGATGAGCCCTCATGTGTCATGCCAGA 2030
Oy 760 aAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyr 780
    ||::: |||
Db 2031 CAACATATGCTACTGTAAA----- 2049
Oy 780 rGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrPheSerAs 800
    ||| :::: |||
Db 2050 -----AAGAAGGCTCAGATCAGACACCTCAATTGAAGCAATATATATGCGACTCC 2102
Oy 800 pasn-----PheAsnHisValProTyrArgLeuArgPheAsnAlaGly 814
    ::::: |||
Db 2103 AAGCAAGCAATGCTGGCTTTGCCAAAAACACATGCGCAAGGTTATGTGTCGCGC 2157

```

RESULT 11
 BC032205 2607 bp mRNA linear HTC 07-AUG-2002
 LOCUS BC032205 Mus musculus, clone IMAGE:5342620, mRNA.
 DEFINITION BC032205
 ACCESSION BC032205.1 GI:21618986
 VERSION
 KEYWORDS
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2607)
 TITLE Direct Submission
 AUTHORS Strausberg, R.
 JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 REMARK NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
 COMMENT Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Galthersburg, Maryland;
 web site: <http://www.nisc.nih.gov/>
 Contact: nisc-mgc@nhgri.nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
 Maduro, Q.T., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,


```

Oy 562 ProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLys 581
    |||:|||||:
Db 1628 GGTGATAAGTTTGGAGAAGCTCAG----- 1651
Oy 582 ProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeuValThr 601
    |||:|||||:
Db 1652 -----AGTTGTAATTCGTGAGTTAAATCTGTGAAA 1681
Oy 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuValGlu 621
    |||:|||||:
Db 1682 AGCCAAAGCTCATCTGTATTGAAA-----AGAAAG 1711
Oy 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
    |||:|||||:
Db 1712 CCCAGAGATCAAAACACATGACAGATT-----TCCAAACCTTAACTCATCTCATAGT 1765
Oy 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlnHisSerThrAsnGly 661
    |||:|||||:
Db 1766 GATAAGCTTCAT-----GGTCACTCT----- 1786
Oy 662 PheAspLeuSerThrArgThrLeuGlnHisGlnIleSerArgSerIleIleGlnAsnGly 681
    |||:|||||:
Db 1786 ----- 1786
Oy 682 GlyTyrAsn-ArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaPr 701
    |||:|||||:
Db 1787 GGTCTCTCAAGAGAGCTCCCATCCGTACAAACTGCAC----- 1826
Oy 701 opTGTGTTTTPGlnAspLeuProValAspPheValAsnGlyLysProSerGlnGluAl 721
    |||:|||||:
Db 1827 -----ACTTGG---TGCAATTCAGTCAAAA--CAGAGCCAGAGAGCTCAGCAGAGC 1875
Oy 721 AlaLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMet-- 740
    |||:|||||:
Db 1876 AGCTGTGGAGAGGCAAGTACCAT-----GTGAAGGATGAGCATGACCATCCAGTCAAG 1929
Oy 740 ----- 740
Db 1930 TGAGCATCTTAAGAGAGAGCATAAACTGAAGCCAAAGAAACCTGATAGAACCTGCAGCC 1989
Oy 741 -ArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAl 760
    |||:|||||:
Db 1990 TTCGCCAAAGAGAGAGCCAGCCAGAGTTTCTCTGGATGAGCTCCATGTTTCATCCAGCA 2049
Oy 760 aasnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyLysAsnAlaTy 780
    |||:|||||:
Db 2050 CAACATAGCTACTGTATAA----- 2068
Oy 780 rGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrSerAs 800
    |||:|||||:
Db 2069 -----AAGAAAGGTCAGATGCAGACACCTCAATTGAAGCAAAATATATGTGACTCC 2121
Oy 800 pAsn-----PheAsnHisValProTyrArgLeuArgPheAlaGly 814
    |||:|||||:
Db 2122 AAGCAAGCAATGTGGGTTTTCAGAAAAAACACATGAGCAAGGTTTATGTGTGCGC 2176
    |||:|||||:
RESULT 12 2837 bp mRNA linear HTC 26-JUN-2002
LOCUS BC032614
DEFINITION Homo sapiens, similar to RAB5 interacting protein 3, clone
IMAGE:5557263, mRNA.
ACCESSION BC032614
VERSION BC032614.1 GI:21595728
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2837)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (06-JUN-2002) National Institutes of Health, Mammalian

```

```

REMARK
COMMENT
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC/DCPD/PTP
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov.
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lalic, P., Legaspi, R.,
Maduro, O.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantford, S., Thomas, P.J., Touchman, J.W.,
Tsurgoun, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAC Plate: 69 Row: m Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gl: 13376344
This clone has the following problem: frame shifted.
FEATURES
source
1. 2837
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /clone="IMAGE:5557263"
   /tissue_type="Skin, melanotic melanoma."
   /clone_id="NIH-MGC_72"
   /lab_host="DH10B"
   /note="Vector: PCMV-SPORT6"
BASE COUNT 581 a 1000 c 805 g 451 t
ORIGIN
Alignment Scores:
Pred. No.: 0.0627 Length: 2837
Score: 117.50 Matches: 143
Percent Similarity: 33.38% Conservative: 111
Best Local Similarity: 18.79% Mismatches: 252
Query Match: 2.49% Indels: 256
DB: 11 Gaps: 35
US-09-914-168-2 (1-919) x BC032614 (1-2837)
Oy 22 ProLeuMetThrSerGlnAlaLeuAlaGlnIleAsnAsnProAlaAsnIleLeuAsnHis 41
    |||:|||||:
Db 1099 CCCATGATGACTCTCGAGAGCTC-----CCATGCCCACTGACAGGCTG 1143
Oy 42 ValProAlaHisAspThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeu 61
    |||:|||||:
Db 1144 GGGCCCTCAGAGAGAGAAAGATGACAGCGGAGGAGCTCCATCCCTTCACAGAGTGC 1203
    |||:|||||:
Oy 62 ThrProGlnGlnIleGlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSer 81
    |||:|||||:
Db 1204 CCCGCCCGCCACTGCTCGAGAAAGAACCTTCC-----ACTGCCCTCCAGAGAGC 1257
    |||:|||||:
Oy 82 GlnAlaLeuAspValValAsnPhaAspAspGlnSerPro---IleSerArgIleGlyGlu 100
    |||:|||||:
Db 1258 CGCGTTTCCGAGAGGAGTCTCTTAGAAGACCAAAAGTCCGGGATGGCGGACAGAGGGAGAC 1317
    |||:|||||:
Oy 101 Gln-----SerProProLeuGlyLeuAspMetSerValIleGlnGluThrThrProLeu 118
    |||:|||||:
Db 1318 CAGCTCAGCCTGCTCCCAAGG-----ACATCAGAGGCGCTGAGAGACACACCC--- 1368
    |||:|||||:
Oy 119 SerLeuGlnGluLeuPheAlaGlnGlnSerThrGlnMetGly-----IleAsn 134
    |||:|||||:

```

Db	1369	-----	-----	CCGAGAGACAGGAGCAAGGCCAGACAGAGGTGAA	1407
QY	135	ProAsnAspTyrIleProGluTyrGlnGlnProAsnSerGluValValPro	154		
Db	1408	GCACGCACAT-----	CCTCACAGCATGCCAGACGTCCAGAGCA	---GCCAAMACAAACC	1458
QY	155	ProThrLeuGluProGluGlyProGlyLeuIleTyrArgLeuTyrAlaArgLeuAsn	174		
Db	1459	CCAGTCCCGCCCCCAGAGAAAAAGGATCTCTCCGACACTGGCTCGACCTC	-----	1512	
QY	175	AspGlyValAsnLysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGly	194		
Db	1513	-----	CCAGCTCCTTAGAGAACCGTGAGCTTGACACAGAGCCGATGGCTTG	1560	
QY	195	GluThr-----	-----	SerAlaIleGlySer	201
Db	1561	GAGCACACCCACGGCGGTCCACCCAGAGAGGCCAAAGCCCTGCTTCTCAGAGCTGGACT	1620		
QY	202	SerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleThrGln	221		
Db	1621	CAGCACCTCTCTCCAGCGCCACCTGCGCAT	-----	TCC	1653
QY	222	GluSerIleMetAspLeuAsnGlnSerIleIleProArgLeuArgGlnThrAlaLeuValAla	241		
Db	1654	CAGAGCTCTCCACAGTTCAAGGCTCC	-----	1680	
QY	242	AlaArgAlaValGlyTyrAspIleAspLeuSerIleIleArgAsnSerIleGlyGlu	261		
Db	1681	-----	CTGGCTCTCTCTCAGACAGCGTTGGG	1707	
QY	262	ValAspValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGlu	281		
Db	1708	GTCGCTCTCATGGCCACCGCCAG	-----	GACTCTTACTCCACGACG	1749
QY	282	ValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeu	301		
Db	1750	AGCAGCAGAGAGAGCTGAGACAGCTTCACACGCCCGAGCTGAAGAAGAGCCCTCATTG	1809		
QY	302	LeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsn	321		
Db	1810	ATCTCGCGCAGAGCTCGGCCCGCGCTCAGCTTTGCTTGCACATTCACAGCATGTTTC	-----	1863	
QY	322	AlaSerIleGluHisGlyTyrPhe---AspGlyArgTyrPheAspArgSerValAspAl	340		
Db	1864	-----	CACGCTTCTCTCTCCAAACCGCAAGCTTACACAGAGGTG	1905	
QY	341	IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyr	359		
Db	1906	-----	CTGGAGCTGGCCACAGCAAGCGCTCGATTGGC	1941	
QY	360	---ArgPheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThr	378		
Db	1942	AGCCTGTGCAGACATCAACAGCTGTACACCTCGGAATGATGGCGCGAG	---ACATCC	1998	
QY	379	AspProAspLysLeuProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMet	398		
Db	1999	AGCAGCAGACATCTCGCAGGAGATTCCACACATGATGACCACACTC	-----	2043	
QY	399	GlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArg	418		
Db	2044	---AAGACCTACTGCTGGAG	-----	2061	
QY	419	TyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnsp	438		
Db	2062	-----	ACACCGGAGCTC	2073	
QY	439	GlnValSerPheGluGlnSerSerSerArgTyrThrGluProAlaGlnValAspLeuSer	458		
Db	2074	-----	AAGGCCGTGGTGACCCGCCCTGCATCCGAGAG	2109	
QY	459	ThrLeuGluProValIleGluThrVal	-----	Glu	468
Db	2110	GAGCTCAACCAATTTAGAGTGTGGCTTGTACAAATGTCTCGAAGCCCTGAGAA	2169		

OY	469	LeuThrAspGlyIleLeuMetAspLysSerProIleGluPheSerAlaSerAsnLeu	468
Db	2170	GCATCATCATATGCTGCCTCATCATGATCCACAGCAAGAGTGGTTCGCTGCACAGCTTCAG	2229
OY	489	GlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTYrAspMetProAsp	508
Db	2230	GAGAACAGTTAGTATGATCTGCGCCACCAACCATGACTA-----	2271
OY	509	ArgValIleuAlaIleAsnHisAspGlyValAsnArgSerIle-----	523
Db	2272	-----GGTGTGACCAACAGCGTGCAGAGTGCCCATG	2304
OY	524	LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer	543
Db	2305	ATGGAAGAATCTCGTCACAACTTCCACAGCATGCACAAAGGCTTACTACCTGAGAGAAG	2364
OY	544	GluAsnGluValIleAsp-----LeuProGluArgThrAlaLeuAlaAsnArg	559
Db	2365	ATCTCATCTCGTCACAAAGCTCATCTACAGACTCCATAGGCCCTCGGCAACCA	2424
OY	560	LysThrProAlaAspValTYrGlnSerIleLysValProLeuTYrValPheValAlaSer	579
Db	2425	GCGAAGCCCATAGGCGGAGTACTCTCGCTGCTCATAGTATGCTGCGCCGAC	2484
OY	580	Asp-----LysProArgAspGlyGlnIle	587
Db	2485	AACCTCAGCGAGATCTCTCATCTGAGAGTACATGATGAGAGCTCATGAGACCCGCCCTG	2544
OY	588	GlyLeuGlyTYrGlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeu	607
Db	2545	CAGCTGGGGGAGGTTCTTACTATCTGACCAACCTACGGGGCCCTGGACAC-----	2598
OY	608	IleAsnArgAspGlyTYrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysIleGly	627
Db	2598	-----	2598
OY	628	ValLysLeuTYrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThr	647
Db	2599	ATCAGACTAC-----GACAAATCATCAGGTG---	2625
OY	648	LeuGlyTYrGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArg	667
Db	2626	-----ACCCGG	2631
OY	668	ThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGlyLysTYrAsnArgThrTYr	687
Db	2632	CAGCTGACTGTGGAGAGTCCAGACTCTCATCCACGCC-----TGGGACGGCGGGCGT	2682
OY	688	SerLeuArgTYrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTYrGlnAsp	707
Db	2683	ACTCTCAC-----AAGGCCGGCGCTCCCGCTCTCCGTACAGGAC	2724
OY	708	---LeuProValAspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyVal	726
Db	2725	TTTCATCTCGTGTCTGTACTCTG-----GAGCCGAGCAGAGGGCGCGAC-GCTGGCGCTC	2777
OY	727	Ala	727
Db	2778	GCG	2780
RESULT 13			
AV557254/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Arabidopsis thaliana			
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
thale cress.			
607 bp mRNA linear EST 06-SEP-2000			
AV557254 Arabidopsis thaliana green siliques Columbia Arabidopsis			
Thailand cDNA clone S006409F 3', mRNA sequence.			
AV557254			
AV557254.1 GI:8728669			


```

melanogaster cDNA clone LD39445 5prime similar to M31617: trx
FBgn0003862 PID:g158818 SWISS-PROT:P20659, mRNA sequence.
A1519730
VERSION A1519730.1 GI:4425584
KEYWORDS EST.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
          Ephygriidae; Drosophilidae; Drosophila.
          1 (bases 1 to 544)
REFERENCE Harvey D., Brockstein P., Hong L., Evans-Holm M., Su C., Tsang G.,
          Lewis S. and Rubin G.M.
          BDCP/HIMI Drosophila EST project
          Unpublished (2001)
TITLE BDCP
JOURNAL Contact: Stapleton, M.
COMMENT Lawrence Berkeley National Lab
          One Cyclotron Rd, Berkeley, CA 94720, USA
          Fax: 510 486 6798
          Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
          Plate: 394 row: D column: 9
          High quality sequence stop: 543.
FEATURES
          location/Qualifiers
          1..544
          /organism="Drosophila melanogaster"
          /db_xref="taxon:7227"
          /clone="LD39445"
          /clone_1ib="LD Drosophila melanogaster embryo port2"
          /sex="male and female"
          /dev_stage="0 to 24 hours mixed stage embryonic"
          /lab_host="XLI Blue"
          /note="Organ: embryo; Vector: PORT2; Site_1: EcoRI; Site_2:
          /XhoI; Sized fractionated cDNAs were directly ligated into
          port2."
          DB: 155 a 161 c 118 g 110 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 0.0155 Length: 544
Score: 112.00 Matches: 39
Percent Similarity: 45.16% Conservative: 31
Best Local Similarity: 25.16% Mismatches: 63
Query Match: 2.37% Indels: 22
          Gaps: 7
US-09-914-168-2 (1-919) x A1519730 (1-544)
OY 13 MetProValAlaLeuAlaIaTyLeuProLeuMetThrSer----- 26
DB 84 ATGCCCAATTATACACTTGTCTGAGCAGCAGTACTGCAATTCGCAATTGTGATGCAACT 143
OY 27 GlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisasp 46
DB 144 CAAAGCTTTGCGACGACGAG---GAGCTGCTAACCGCGTGCAGCATTTCTCCACAAGCAGC 200
OY 47 ThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGluGlnIle 66
DB 201 AGCAGT---AGCAGTACCAACTGCTCACTGCCACCAATGTGCTTAATCTCATGCAACAA 257
OY 67 GlnAlaIArgLeuAsnAlaIaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspVal 86
DB 258 CAAGCACCATTCAACCAACCACTGCTCCACAGCAGATGTCCTCCAGTTTGTACCCAGATG 317
OY 87 ValAsnPhaSpaSpGlnSerProIleSerArg-----IleGlyGlnGlnSer 102
DB 318 CACACAGCGCCAGGAACTGCTCCCTTGTCCACAGCAGTGTCCCGTAGTTTGTACCCACT 377
OY 103 ProProLeuGlyLeuAspMetSerValIleGluGlnThrThrProLeuSerLeuGlu 122
DB 378 CCGCCGAACCTGTTGAACAGCCATATAATCAATCAATGACTAGCGGACAGTGTCCAG 437
OY 123 LeuPhaIaGlnGlnIleSerThrGluMetGlyIleAsnProAsnAspTyrlleProGluTyr 142

```

```

DB 438 TGCTATGCCCCCAAAAGTCAACA-----CTACCATCA----- 476
OY 143 GlnGlyGlnGlnProAsnSerGluVal-----ValValPro 154
DB 477 GAGCGGAACTCAAAAGTTAGTTTCAGTTTGGAGACATTTGCCA 521
RESULT 16
LOCUS BC025329 3144 bp mRNA linear HTC 07-AUG-2002
DEFINITION Homo sapiens, clone IMAGE:4111278, mRNA.
ACCESSION BC025329
VERSION BC025329.1 GI:22137580
KEYWORDS HTC.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
          1 (bases 1 to 3144)
REFERENCE Strausberg R.
          Direct Submission
          Submitted (05-MAR-2002) National Institutes of Health, Mammalian
          Gene Collection (MGC), Cancer Genomics Office, National Cancer
          Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550,
          USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
          Email: gcaps-r@mail.nih.gov
          Tissue Procurement: ATCC
          CDNA Library Preparation: Rubin Laboratory
          CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Genome Sequence Centre,
          BC Cancer Agency, Vancouver, BC, Canada
          Info@cgsc.bc.ca
          Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
          Susana Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
          Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
          Sen Lee, Victor Ling, Carrie Mathewson, Candice McElvay, Steven
          Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
          Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stolt,
          Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
          George Yang, Scott Zuyderduyn, Marco Marra.
          Clone distribution: MGC clone distribution information can be found
          through the I.M.A.G.E. Consortium/LNL at: http://Image.Lnl.gov
          Series: IRAL Plate: 42 Row: C Column: 5
          This clone was selected for full length sequencing because it
          passed the following selection criteria: Genomescan gene prediction
          This clone has the following problem: frame shifted.
FEATURES
          Location/Qualifiers
          1..3144
          /organism="Homo sapiens"
          /db_xref="taxon:9606"
          /clone="IMAGE:4111278"
          /tissue_type="muscle, rhabdomyosarcoma"
          /clone_1ib="NIH-MGC_17"
          /lab_host="DH10B-R"
          /note="Vector: pOT7"
          DB: 952 a 696 c 759 g 737 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 0.457 Length: 3144
Score: 111.00 Matches: 90
Percent Similarity: 34.77% Conservative: 55
Best Local Similarity: 21.58% Mismatches: 153
Query Match: 2.35% Indels: 119
          Gaps: 17
US-09-914-168-2 (1-919) x BC025329 (1-3144)
OY 445 SerSerSerSerArgThrGluProAlaGlnValAspGlnSer-ThrLeuGluProValIle 464
          :::::

```

D	1005	ACTCTCGTCGACCCGGAACAAGAGCTGGACGAATAATTAAAGATTCACATTCACAGAGCAGGANTT	1064
Q	464	egLurth---ValGLuLeuthrAspLyIle-----LeuMetAspIleSerProI	480
D	1065	GGAAGTCTCTTGAGGCTTACAGAGTGAATTTGGGTTCGTCCATTAATTAAGAGTCCAGA	1124
Q	480	egLurPheSerAlaSerAsnLeuIleGLInAspLysLeuAsnLeuValAlaAlaLysAlaAr	500
D	1125	AGAAGCTGGCTTTGGCAAGCTTTTCACAGAAATCAATATGGCATCTATACACAGCTGGACACCAG	1184
Q	500	gHISLeuTyraSPMePProAspAspArgValLeuAlaAlaLeuAsnHISAspAspGLyValAs	520
D	1185	TCNATTCCTTGAA-----CTCACAGAA---CAGAGATGACAGAGTGGCA	1223
Q	520	narGSerIleLeuGLyArGLIeserAspAlaValSerAlaValAlaArgAlaIleLeuPr	540
D	1224	CAATCAGCTATTGGTTCACCCAGCAGCCCTTCATGCTCTGTATGGACT-----CC	1271
Q	540	oAspGLuSerGLuAsnGLuValIleAsp-----	549
D	1272	CTGAGCAATGAGAATAGGTCTATCGACAGACGGGCCAGCTGTGCAAGTCAACGCAAGTGA	1331
Q	550	-LeuProGLuArGLIleAlaLeuAlaAsn-----ArgLy	560
D	1332	GTTCCCCAGTAAGAGGTCACTGACCAACATCTTCACAAAGCATCTGATGATCACAGCTCAAG	1391
Q	560	sThrProAlaAspValTyGLInSerLysLysValProLeuTyValPheValAlaSerAs	580
D	1392	ATTCAcAGTGCAAATTGAGAGAAACTGCTCCCTCAAGCTGCTAAAGTTTC-----	1440
Q	580	PLysProArgAspGLyGLInleGLyLeuGLYtrProLysArgAspThrGLYtrAlaGlyLeu	600
D	1441	-----TTTGGCTACGATGCACAGACGAATCAGAGCT	1469
Q	600	lThrLysPheGLuHISAsnLeuIleAsnArgAspGLYtrGlnAlaGLyAlaGLuLeuAr	620
D	1470	GGAATAATATGATGAAAACTCATATAAAAGACAGCTGACAGAGTGGACACCAACATTCG	1529
Q	620	g-----LeuSerGLuAspLysLysGLyValLysLeuTyAlaThrLysPr	635
D	1530	ATACTACTTTGAAAATATCTCAAAATTCAGCATTCCTCGATGACAGTAAcGtGtGTCCACTC	1589
Q	635	oLeuSerHISProLeuAsn---AspGLInLeuArgAlaThrLeuGLYtrGlnGLInGLuVa	654
D	1590	CAACACAGCTCCCATGTGATTTTAAGGCCCTTAATAAAACACACTTGGGGTTTCCTTGATACG	1649
Q	654	lPheGLYHISSerThrAsnGLYrPheAspLeuSerSerThrArgThrLeuGLInHISGLuIleSe	674
D	1650	GTTTGAAGACGCTGTGATTAATCAATCATTCACACTCGGGTACATCCCTATGAGACCA	1709
Q	674	rArgSerIleIleGLInAsnGLYtrPAsnArgThrTySerLeuArgTyArgGlyLeuAs	694
D	1710	GGAGTTCATCATCATCAT-----	1725
Q	694	PLysLeuLysThrGLInAlaProProGLuThrTrpGLInAspLeuProValAspPheValAs	714
D	1726	-----GATATCCTC-----GATATCCTC-----	1734
Q	714	nGLYAsProSerGLInGLuAlaLeuLeuAlaGLyAlaValaIleHISLysThrValAlaAs	734
D	1735	----AAACATTTCCAGAGCACTCCTCATGCCAGCAGCGAGCTGCATATCCGGGATCAGTGA	1790
Q	734	PaSnLeuValAsnProMetArgGLYtrArgGLInArgTySerLeuGLuValaGLYserSe	754
D	1791	TTTTCTTGGCAATCCATG-----	1809
Q	754	rgLYLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGLYIleSerGLYValTyrSe	774
D	1810	-GGGCTTTGAATGATGATTTCT-----GAAGGGGTATACGAGCTGATAAA	1853
Q	774	rPheGLYAspAsnAlaIaTyGLY-----SerAsnArgAl	785
D	1854	ATATAGGA---AAATGCGGGGGCTCATCGAAATGTTACACACGGAGTAACTGATTCG	1910

```

Ory 14 ProValAlaLeuAlaIaIyIreUpProLeuKetHrSerGlnAlaLeuAlaGlnGlnAsn 33
      |||||   :::   |||||   |||||   |||||
Db   5 CCACTACTGCATTCGCAATT-----GTATGAGACCTCAGCTTGGAGCAGCAG--- 55

```



```

QY 203 ---HisGlnLysThrGlu-----ProTyrAlaAsnIleLysAlaIleGlu 217
      |||||  |||
Db 824 ATACACACATGAGAGAGAACACACACAGTACAGAGAGACTCCAGGCAACCCCTGCCAA 883
QY 218 AspIleThr-----GlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArg 235
      |||||  |||||
Db 884 GAGCTAGCTGATTACAGACAGATTACCCAGACACTGATAGTGAACGAAGGCTTGGG 943
QY 236 GlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIle 255
      |||||  |||||
Db 944 GAAGAGAGAGGTTATT-----CTGATG 964
QY 256 ArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGlyLuproValTyrIle 275
      |||||  |||||
Db 965 GAGCTCTTTATGTCAGACAGACCATAGTGGACACTTATGTCAGAC-----ATT 1015
QY 276 AspTyrArgAlaValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrVal 295
      |||||  |||||
Db 1016 GAATAC-----TTCCGCTCTCTT 1033
QY 296 AlaAspGluValProLeuLeuIleGlyAspValPheHisGlyLysTyrGluThrLys 315
      |||||  |||||
Db 1034 CTAGATGAG-----CATCACATTTCTTATGTCATAGAT 1066
QY 316 LysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPleuAsp 335
      |||||  |||||
Db 1067 GAAGATGTA-----AAAGTGGCGCTATATGCAATTAGACAGACCTTATCATGAC 1117
QY 336 ArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThr 355
      |||||  |||||
Db 1118 -----CTCCGCTGAGATGCC----- 1132
QY 356 GlyThrGlnTyrArgPheAspGluValValPhePheThrIleAspProLysThrAsnGln 375
      |||||  |||||
Db 1133 -----CGTTTGACGG----- 1144
QY 376 LeuThrThrAspProAspLysLeuProValLysArgGluLeuGluGluLeu--- 394
      |||||  |||||
Db 1145 -----GAGCAGCTTCTTGGT 1159
QY 395 -----ThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaVal 407
      |||||  |||||
Db 1160 GTCCACACACATTTAACAATCTTTGAAATGCGCAACAGACATAGAGAACTCAA 1219
QY 408 ArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIle 427
      |||||  |||||
Db 1220 GAATGATAGCG-----GCACCTCAAGAACGACATCCCATATG-----GAGCGAATT 1267
QY 428 ValPheProGluArgGlu----- 433
      |||||  |||||
Db 1268 ATTGAATCTGAGCAGAAAGAAAGAACAGCGCTTGACGCCACTTAGAGATCAAGGCC 1327
QY 434 GlnIleGlnAsnAspGlnValSerPheGlu-----GlnSerSerSerArg 449
      |||||  |||||
Db 1328 ACAGTGGCCAGTGCAGATAGATGAATCGCCTGAGAGCTGCGAGATGAAG 1387
QY 450 ThrGluProAlaGlnVal-----AspLeuSerThrLeuGluPro 462
      |||||  |||||
Db 1388 CAGAAAGTGGCAGAGCTGATTCTATTCATTACTCTGAGCAAAATCTGATTATTCAGGAC 1447
QY 463 ValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPhe 482
      |||||  |||||
Db 1448 CTCCTGAGAGTGTGAGCGG-----GACAAAGAAAGACAGAGACT 1489
QY 483 SerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeu 502
      |||||  |||||
Db 1490 TTGGCTAGTAGCTTGCAGGAGAT----- 1513
QY 503 TyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSer 522
      |||||  |||||
Db 1514 -----CTGGCTCATACCCGGAATGATGCG----- 1537

```

```

QY 523 IleLeuGlyArgIleSerAspAlaValSerAlaValAla-----ArgAlaIleLeu 539
      |||||  |||||
Db 1538 -----AATCGATTACAGATGCCATTGCTAGAGTAGAGATCAATACGAGCG-----TTC 1588
QY 540 ProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArg 559
      |||||  |||||
Db 1589 CAAGAAGACGCTAGAAACAAATTCAGGATTTG-----AAATAGACGTTAGAA 1636
QY 560 LysThrProAlaAspValTyrGlnSerLysLys 570
      |||||  |||||
Db 1637 AAATTAAGATCAGACCTGGATGAAAAA 1669

RESULT 19
AY104509 1501 bp mRNA linear HTC 25-MAY-2002
LOCUS AY104509
DEFINITION Zea mays PCO104742 mRNA sequence.
ACCESSION AY104509
VERSION AY104509.1 GI:21207587
KEYWORDS HTC.
SOURCE Zea mays.
ORGANISM Zea mays.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
Clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 1501)
Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitesitt,M.S.,
Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.
Maize Mapping Project/DuPont Consensus Sequences for Design of
Overgo Probes
Unpublished (2002)
2 (bases 1 to 1501)
Coe,E.C.
REFERENCE
JOURNAL Direct Submission
AUTHORS Submitted (25-Apr-2002) Maize Mapping Project, University of
TITLE Missouri, Columbia, MO 65211, USA
JOURNAL Location/Qualifiers
FEATURES
source
1..1501
/organism="Zea mays"
/db_xref="taxon:4577"
/clone="PCO104742"
/clone_lib="Maize Mapping Project/DuPont Consensus
Library"
/note="this sequence is part of a project of EST
assemblies resulting from the application of public
contigs to seed DuPont contigs; this resource was
assembled by DuPont as part of a collaboration for the
overgo addressing of BACs in conjunction with the Maize
Mapping Project"
BASE COUNT 308 a 489 c 418 g 286 t
ORIGIN
Alignment Scores:
Pred. No.: 0.215 Length: 1501
Score: 109.00 Matches: 98
Percent Similarity: 31.26% Conservative: 53
Best Local Similarity: 20.29% Mismatches: 168
Query Match: 2.31% Indels: 164
DB: 11 Gaps: 23

US-09-914-168-2 (1-919) x AY104509 (1-1501)
QY 439 GlnValSerPheGlnGlnSerSerSerArgThrGluProAlaGlnValAspGluSer 458
      |||||  |||||
Db 149 GAACCTATCTCGCTCGCGGCTTACAGTCACTCAATTTGGCAACAGTTCGAAGC 208
QY 459 ThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSer 478
      |||||  |||||
Db 209 AATACCGACCA-----AGGACGACGACGAGCGTGCATGGCGACGCTCTCC 253
QY 479 ProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLys 498
      |||||  |||||
Db 254 -----TCTGACACCGCGCTGAGTAGCGGC-----ACCGGCGCGCGG 289

```


DB:	11	gaps:	35
US-09-914-168-2 (1-919) x BC022490 (1-2083)			
OY 311	LysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAsp	330	
Db 54	AAAAAGCGAAGAAAGAACACTTTCATTATGACGACCCAGC-----CATGCGACAGCTAGCA	107	
OY 331	GlyArgTyrPheAsnAspArg-----SerValAspValIleLeuProAspAsnThr	346	
Db 108	GCCGCTGCTTTCCACAGCGGACGACAGCTCGGAGCTCGGACGTGTGTTTCCCTCAAGATTG	167	
OY 347	AlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhe	366	
Db 168	CTAAGCTGCTGCTTTATT-----ACTGAAAGAAAGATGTGGCGAGATTGTTTTC	215	
OY 367	PheThrIleAspProLysThrAsnGluLeuThrThrAspProAspLysLeuProValLys	386	
Db 216	TTTACTCTGAC-----	227	
OY 387	ArgGluLeuLeuGluGluLeuLeuThrValAsnMetGlyAlaTyrAsnLeuGlnAla	406	
Db 228	-----TGATCACTTGTCTTGGCCGACAGCTATAC-----	257	
OY 407	ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMet-----	422	
Db 258	AACTTTCGAAGACGATGACGAGCATAGAAAGAACAAATACAGTCCAGCATGGCTCC	317	
OY 423	ValAsnThrGluIleValPheProGluArgGluGluIleGlnAsnAspIleValSerPhe	442	
Db 318	TGCAGCTCACTTCTCTCTCTCCACAGATGACAGCACTGCCGCTTCTCTCGAGCCCTAC	377	
OY 443	GluGlnSerSerSerSerArgTyrGluProAlaGlnValAspGluSer-----	458	
Db 378	GTGTCCATTCGTGTGTCAGAGGAGGAGCCGCCCTCGAATAGATGACTCGGTGCAGAGGCTG	437	
OY 459	---ThrLeuGluProValIleGluThr-----ValGlu-----	468	
Db 438	CAAGTCTGTGGAAACATCATGGAACAAACACACTAGTGCTAATGAAGCTTGGAATATAT	497	
OY 469	LeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIle	488	
Db 498	ATCCAGGACACACATGTAAGAAAGAAATGTTACAGATTACAG-----CAGAAATGCAGTA	548	
OY 489	GlnAspLys-----LeuAsnLeuValAlaAla	497	
Db 549	CAGGACGACAGCGGCTGTGATGATGAATAAGACAAACACTGTTGAACCAACAGCGGAG	608	
OY 498	LysAlaArgHisLeuTyrAspMetProAspAspArgValLeu-----	511	
Db 609	CAAACGCGGAAGTTAATCATGATGTG---GAAACCCCAAGTATTTAATCATGACCAGCAGACTT	665	
OY 512	---AlaIleAsnHisAspAspGlyValAlaSerG-----	521	
Db 666	GAACCTTCAGCTCTTGGAACACTCCCTCTCGACAAACAAATTTGGAAAAACAGATTTGGAC	725	
OY 532	---SerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeu	539	
Db 726	CAGACCACTGAATTAACAAATTTGCAAGATTAAGAACAGTTTCCTAGAAAGAAAGAGTGCTA	785	
OY 540	ProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArg	559	
Db 786	GCTATGGAA---GACCAACACATCATCTCAACATA-----CAGTCATATAAAGAGAG	833	
OY 560	LysThrProAlaAspValTyrGlnSerTyrLysValProLeuTyrValPheValAlaSer	579	
Db 834	AAAGATCAGCTACAGCTGTAGTATCCAAAGCAAAATTCATCATAT-----	878	
OY 580	AspLysProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGlyThrArgLeu	599	
Db 879	-----GAAAGAACTAGAAAAAAATA	899	
OY 600	ValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeu	619	

Accession	Gene	Species	Length (bp)	EST ID	EST Date
D6	GTGACTGCC	-----ACGGTGAATATATTCAGTTCTTCACAGACACATCATAT	947		
QY	ArgLeuSerGIuAspLyIsGlyValIysLeuTYrAlaThrIysProLeuSerHisPro	639			
D6	CTCATGAGACAGATTAAATAACTTACTGACTATGATGTCACATCACTAACTTAAGACACC	1007			
QY	LeuAsnAspGIuLeuArgAlaThrLeuGIyTYrGIuIn	652			
D6	1008	-----ACGTGTTCTAAAGAGAACAAATACGCTTCAGAGACTGT	1046		
QY	653	---GIuValPhe-----GIyHisSerThrAsnGly-----PheAspLeu	664		
D6	1047	GCTGAGATTTCAAATCAGACAGACACACAGAAATGGCATCTACACGTTACATTTCCCTTAAT	1106		
QY	665	SerThrArgThrIleGluHisGIuIleSerArgSerIleIleGluAsnGIyTYrAsn	684		
D6	1107	TCTACAGAAAGATTCACAGCCCTACTGTACATAGAAA---CGTGGAGAGACGGCGGTG---	1160		
QY	685	ArgThrTYrSerLeuArgTYrArgLeuAspIysLeuLYrThrGluAlaProGluThr	704		
D6	1161	---ACAAATTAATTCACGACGAGGAGATGGACGCCGTGTGATTTTCTG-----AGGACT	1208		
QY	705	TYrGIuAspLeuProValAspPheAlaIasnGIyIysProSerGIuGluAlaLeuLeuAla	724		
D6	1209	TGGAAAGATTAATAAGTGGCATTT-----GATAACCTTCACAGAAATATTGGCTGGGA	1262		
QY	725	GIyValAlaValAlaHisIysThrValAlaIAspAsnLeuValAsnProMetArgLYrTYrArg	744		
D6	1263	AATGAGTTTGT-----TCGCACACTGACTAAT-----CAG	1292		
QY	745	GIuArgTYrSerLeuGIuVal-----	751		
D6	1293	CACGCGTATGTGCTTAAATACACCTTAAGACTGTGGAAGGGAATGAGGCTTACTCATTTG	1352		
QY	751	-----	751		
D6	1353	TATGACAACTTTCTATCTCTCAAGTGAAGAACTCAATTATAGATTCACCTTAAGAGACTT	1412		
QY	752	---GIYSerSerGIyLeuValSerAspAlaIAsnMetAlaIleAlaArgAlaGIyIleSer	770		
D6	1413	ACAGGACACAGCGCGCAAAATAGC-----ACATCAGCCCAACAGAAATGAT	1460		
QY	771	GIyValTYrSerPheGIyAspAsnAlaTYrGIySerAsnAlaGluAlaHisGluMet---Thr	789		
D6	1461	TTTACACAAAGGATGAGACACCAACAATGATTGTGCAAAATGTTCACAAATGCTTACAA	1520		
QY	790	GIyGIyIleGluAlaGIyTYrIlePheSerAsp-----AspPheAsnHis	804		
D6	1521	GGAGCC-----TGTGTGTTTGTATGTCATGTGTCCTTCCAACTTGAACGGA	1565		
QY	805	ValProTYr-----	807		
D6	1566	ATGTACTATCCACAGAGAGCAGAACACAAATATAGTTCAAGGCACTTAATGTACTACTAG	1625		
QY	808	---ArgLeuArgPhePheAlaGIyIysAspGIuSerIleArgGIyTYrAlaHisAsp	825		
D6	1626	GAAAGGCTCAGGCTATTCGCT-----CAAGGCCACAAACCATGAT	1664		

AUTHORS Walbot, V.
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University
JOURNAL Unpublished (1999)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 8221
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 614073 row: B column: 08.
Location/Qualifiers
1..588
/organism="Zea mays"
/cultivar="W23"
/db_xref="taxon:4577"
/clone_lib="614 - root cDNA library from Walbot Lab"
/tissue_type="root"
/dev_stage="3-4 days old"
/lab_host="XLOLR"
/note="Organ: root; Vector: pBluescriptII SK+; Site: 1; EcoRI; Site: 2; XhoI; 3-4 days old root tissue from Walbot Lab (Lm)"
BASE COUNT 182 a 156 c 112 g 138 t
ORIGIN

Alignment Scores:
Pred. No.: 0.0939 Length: 588
Score: 106.00 Matches: 34
Percent Similarity: 41.60% Conservative: 18
Best Local Similarity: 27.20% Mismatches: 41
Query Match: 2.24% Indels: 32
DB: 10 Gaps: 5

US-09-914-168-2 (1-919) x AM053211 (1-588)

OY 788 MetThGlyGlyIleGlnAclYrIleTrrPserAspAsnPhaSnHlSValProTYR 807
:::|||||
DB 543 TTGACAGCGCGATCCATGTTGTGGC-----GACATGGCACCCTTAT 505
:::|||||

OY 808 ArgLeuArgPhePheIleGlyAspGlnSerIleArgGlyTrrAlaHisAspSerLeu 827
:::|||||
DB 504 CAA--GCTTTGGCCATGAGTGGCTGAGTCCGAGGCTATGCT----- 460
:::|||||

OY 828 SerProIleSerAspGlyTrrLeuThrGlyGlnValLeuAlaValGlyThrAla 847
:::|||||
DB 459 -----GAGGGCGCTGTTGCTTCAGAAAGACTATGCTAACTTCTAAGTGC 415
:::|||||

OY 848 GluTrrAsnTrrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsn 867
|||||
DB 414 GAATATACGATTCCTCTTGGTGAAGCATCTGGAGGTTCTATTTCATGTGACTGTGGAGT 355
|||||

OY 868 AlaTrrAspLysGlyPheThrAsnAspThrLys----- 878
|||||

DB 354 -----GACTTAGGGTCTGCTGCTCATGTCCCTGTAATCCACCTCGCGGAGGGGAAA 301
:::|||||

OY 879 -----IleGlyAlaGlyValGlyValArgTrrPalaSerProValGlyGlnVal 894
|||||

DB 300 CCAGGATTTGCTATGATTTGGATACGAGACTTCAACACTGACCTAGGTGACAGATC 241
|||||

OY 895 ArgValAspValAla 899
|||||

DB 240 CGTGTGACATACGCT 226
|||||

RESULT 22
LOCUS BE788630 758 bp mRNA linear EST 20-OCT-2000
DEFINITION 601475986F1 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:3878925 5',
mRNA sequence.
ACCESSION BE788630
VERSION BE788630.1 GI:10209828
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 758)
NIH-MGC <http://mgs.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgs@bbs-rcmail.nih.gov
Tissue Procurement: DCTD/DRP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
Plate: LLM9643 row: k column: 22
High quality sequence stop: 339.
Location/Qualifiers
1..758
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:3878925"
/clone_lib="NIH_MGC_68"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pCMV-Sport6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.8 kb. Library constructed by life
technologies." 12
DB: 12 Gaps: 8

Alignment Scores:
Pred. No.: 0.194 Length: 758
Score: 105.00 Matches: 53
Percent Similarity: 42.64% Conservative: 31
Best Local Similarity: 26.90% Mismatches: 85
Query Match: 2.22% Indels: 29
DB: 12 Gaps: 8

US-09-914-168-2 (1-919) x BE788630 (1-758)

OY 521 ArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuPro 540
|||||
DB 59 CGGAAAGTGTCTC-----ATCAGTGACAGCGCTTGACCTTGTCCGGAAGATCTTGCCA 112
|||||

OY 541 AspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLys 560
|||||
DB 113 GATGGAGGG-----CTGCAGGTGTGTGGAAGAACGAAACCTTGCAAAAGAGAG 160
|||||

OY 561 ThrProAlaAsp-----ValTrrGlnSerLysLysValPro 572
|||||

DB 161 CTGATAGCGGAGCTGAGAGCTGTGAAGGCTTATTTGCTGCTCCACCAAGGTGACC 220
|||||

OY 573 LeuTrrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGly----- 590
|||||

DB 221 GCTGATGTCATCAACGACGTGGAACCTCCAGGTGTGGGACGGGCTGGCAGAGTG 280
|||||

OY 591 -----TrrGlySerAspThrGlyThrArgLeuValThrLysPheGluHis 605
|||||

DB 281 TGGACATGTGTGATCTGAGGCGCAACAGGA---AGGCACCTCTATGTATGAACAC 337
|||||

OY 606 AsnLeuIleAsnArgAspGlyTrrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLys 625
|||||

DB 338 CCCCAACAGGAGGAGACAGCCCTCAAGATGGCCGCAACCAACATCAACATAGACT 397
|||||

OY 626 LysGly-----ValLysLeuTrrAlaThrLysProLeuSerHisProLeuAsn 641
|||||

DB 398 CGGGGCAATGAATCACTAGTGCAGCCATCGGCGCACAGAGCAGG-ATGTATCCCGCGGAC 456
|||||

OY 642 AspGlnLeuArgAlaThrLeuGlyTrrGlnGlnGluValPheGlyHisSerThrAsnGly 661
|||||

Db 457 GGGCGACAGAGGCGACTGTCGATAGACAGAGGACCGGACCGACAAA---AATGGG 513
Oy 662 pheaspleuSerThrArgThrleuLuhSgluIleSerArgSerIlellelelnsnngly 681
Db 514 AGAACCCGACGACGACCAACCCGATV---GCCACTCGGGGGCGCAACGACGAG 570
Oy 682 GlyTPAsnArgThrTySerleuArgTyArgleuAspLysleuLysThr 698
Db 571 AAGGAGTCTAGACAACTCGGTGACACACGACGACACCCCTCGGAGC 621

RESULT 23
BH770729/c 1557 bp DNA linear GSS 01-MAY-2002
DEFINITION LMGtag478 MG1363 Random Sequence Tag Library Lactococcus lactis
subsp. cremoris genomic, DNA sequence.
ACCESSION BH770729
VERSION BH770729.1 GI:20373686
KEYWORDS GSS.
SOURCE Lactococcus lactis subsp. cremoris.
ORGANISM Lactococcus lactis subsp. cremoris
Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
Lactococcus.
REFERENCE 1 (bases 1 to 1557)
AUTHORS Bojotin,A., Ehrlich,S.D. and Sorokin,A.
TITLE Studies of genomes of dairy bacteria Lactococcus lactis
JOURNAL Sci. Aliments, (2002) In press
COMMENT Contact: Sorokin A
Genetique Microbienn
INRA
CRJ INRA, Domaine de Vilvert, 78352 Jouy en Josas cedex, France
Tel: 33 1 34 65 25 16
Fax: 33 1 34 65 25 21
Email: sorokine@jouy.inra.fr
best homologue in strain IL1403 is ywfg (76%)
Class: shogun
High quality sequence start: 30
High quality sequence stop: 1527.
Location/Qualifiers
1. 1557
/organism="Lactococcus lactis subsp. cremoris"
/strain="MG1363"
/db_xref="taxon:1359"
/clone_lib="MG1363 Random Sequence Tag Library"
/note="Vector: pSGM02; Site 1: SmaI; Library of
chromosomal fragments of L.lactis strain MG1363 was
prepared by partial AluI digestion or by sonication."
BASE COUNT 403 a 318 c 253 g 582 t
ORIGIN
Alignment Scores:
Pred. No.: 0.696 Length: 1557
Score: 105.00 Matches: 135
Percent Similarity: 29.24% Conserved: 70
Best Local Similarity: 19.26% Mismatches: 200
Query Match: 2.22% Indels: 296
DB: 17 Gaps: 34
US-09-914-168-2 (1-919) x BH770729 (1-1557)
Oy 181 ProArgleuLysAlaLysPheTyrglnSerGlnSerGlyGluThrSerAlaIleGly 200
Db 1544 CCCCCGTTG-----CTCCGTCAGTAACAACANGTCGAGACATTACACAGCTTAC 1494
Oy 201 SerSerHisGlnLysThrLupProTyArgAlaAsnIleLysAlaAlaLeuLysPleThr 220
Db 1493 ACGACAGACCAAAAATATTTCCA-----GGATATACT 1461
Oy 221 GlnGluSerAlaMetAspLysnnglySerIleProArgleuArgGlnThrAlaLeuVal 240
Db 1460 TATTCATCATCA-----AATGGA----- 1443
Oy 241 AlaAlaArgAlaValGlyTyTyArgPleAspLysPleuSerIle-----IleArg 256

Db 1442 GCAGCACAAGCGAGAAACTATACGTTAATGATCAAACTGTAATATGTTATACAGA 1383
Oy 257 AsnSerIleGlyGluValAlaSpValIlelleHisAspLeuGlyGluProValTyThrLeu 276
Db 1382 AATCAA---GGAGCATGATGTA-----ACCTACATTTGAT 1350
Oy 277 TyrArgAlaValAlaGluValArgGlyLys----- 285
Db 1349 CAGACACAGGCGCAAGTTTGGACGAAGAAGATTGTCGTCGTAAGTATGACCTCGG 1290
Oy 286 -----GlyAlaAsp-----AspLysAlaPheThrValAla 296
Db 1289 AATTATACAAACAGTCGATATTAATCTTACACGATAAAGATATGATTAATGTTCT 1230
Oy 297 AspGluValProLeuLeuIleGlyAspValPhe-----HisIleGlyLys 311
Db 1229 GATGATTATCCAGTCAA---GGACGAGTATTTACGGATACCTCCAGCATTTATGAGTT 1173
Oy 312 TyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluIleGlyTyPheAspGly 331
Db 1172 TATTTAAGCAAAATACAGTTGTAAGTTGGAACAAACAA----- 1131
Oy 332 ArgTyPLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeu 351
Db 1130 -----GTGAATGAAACCATTCATTAT 1110
Oy 352 IleTyArgPThrGlyThrGlnTyArgPheAsp-----GluValAlaPhePheThr 368
Db 1109 GTCTATGAGATGAGACATGAGCTGCAGAGACTATTAAGCAGACCCGCTGAATTTACA 1050
Oy 369 IleAspProLysThrAsnGlnLeuThrTyArgProAspLysLeuProValLysArgLys 388
Db 1049 AAGACCGTACAAACGATATACAGTACGGGTGAAAA----- 1014
Oy 389 LeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyArgLysGlnAlaValArg 408
Db 1013 -----ACATACGAGCAATGAGCAACCGCAAAAT----- 987
Oy 409 AlaLeuSerAsnAspLeuLeuAlaThrArgTyPheAsnMetValAsnThrGluIleVal 428
Db 986 -----GTTAATAGTTT-----CTTAAGTAGTT 963
Oy 429 PheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSer 448
Db 962 TCACCGATTATCAAGAGTAAACACCAGACCA----- 930
Oy 449 ArgTyThrProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGlu 468
Db 929 -----GCTCAATGATGATAATACG----- 909
Oy 469 LeuThrAspGlyIleLeuMetAspLysPheProIleGluPheSerAlaSerAsnLeuIle 488
Db 908 -----GGTGTATACAGCAGATACGACGATATTGAG----- 879
Oy 489 GlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyArgPmetProAspAsp 508
Db 878 -----AAA 876
Oy 509 ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSer 528
Db 875 AAGATATCTATTAACGATATCAAGAGACGATGATGATGATCTATATTT----- 828
Oy 529 AspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluIle 548
Db 827 -----GATGAACGACAGTCAAGTTCTG 804
Oy 549 -----AspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspVal 565
Db 803 ACTAAAAAAGATTGTCGTGCGTGAACGATGATCGCTCAATATATCAACAGGTGCGAT 744
Oy 566 TyrGlnSerLysLysValProLeuTyThrValPheValAlaSerAspLysProArgAspGly 585
Db ----- 111

```

Db 743 ATTAAATCTTACACAGATAAAGCTTACGAATTAGTTTCAAGTGATTATCCAGATGGCGG 694
Oy 586 GlnIleGlyLeuGlyTrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHis 605
Db 683 ACATC-----TTTACTGATGAAGCACAACATTATACCGTGAAATTGAAACAT 636
Oy 606 AsnLeuLeuAsnArgAspGlyTyrGlnIleGlyValGluLeuArgLeuSerGluAspLys 625
Db 635 GGATTAACAAT-----GAACTGATTA 612
Oy 626 LysGlyValLysLeu-----TyrAlaThrLysProLeuSerHisProLeuAsnAsp 642
Db 611 AAAGCTGTAATCAAGTATTCATTATGTTTATTAAGATTAAGTGAAGTAAGCAGCAGAG 552
Oy 643 GlnLeuArgAlaThrLeuGlyTyrGlnGlnGlnValPheGlyLysSerThrAsnGlyPhe 662
Db 551 GACCATACGCTACTCTTGATTTAGTCACAATA----- 516
Oy 663 AspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGlyGly 682
Db 515 -----AGTACTGACAACAGTAACAAAAG----- 492
Oy 683 TrpAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProPro 702
Db 491 -----AAACTTATAGC----- 480
Oy 703 GluThrTrpGlnAspLeuProValAspPheValAsnGlyLysProSerGlnAlaLeu 722
Db 479 -----GACTGC----- 474
Oy 723 LeuAlaGlyValAlaValHisLysThrValAlaAspAsn-----Leu 736
Db 473 -----GGCGCTGACAATGCTGACAGCTTTCCGAAGTC 441
Oy 737 ValAsnPro---MetArgGlyTyrArgGlnArgTyrSerLeuGlnValGlySerSerGly 755
Db 440 GTTCCGCACTGCTCAAGAGTTTATACAGCTGACCAACTTAAGTCTGTGAATGACAGCA 381
Oy 756 LeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyLysSerGlyValTyrSerPhe 775
Db 380 ATCATTCTTCACACA-----GCAGATATTCTCAGTAACGCGTTTCTTAT 339
Oy 776 GlyAspAsnAlaIleTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGly 795
Db 338 ACAAGAAAT-----CAAGCAACGATTGATGTCGACT 309
Oy 796 TyrIleTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGly 815
Db 308 TATATT-----CATGAACCACTGGGAAGTGTATCTCAAAAGATTGTCAGCGCGA 255
Oy 816 AspGlnSerIleArgGlyTyrAlaHis---AspSerLeuSerProIleSerAspLysGly 834
Db 254 ACTGGAGATTCTTCAACAGCTATACACAGCGTTGATTAATCTTACACAGATAAGGCT 195
Oy 835 Tyr 835
Db 194 TAC 192

```

```

RESULT 24
CNS071R5/ 1021 bp DNA linear GSS 07-JUL-2001
LOCUS 1021 bp DNA linear GSS 07-JUL-2001
DEFINITION Kluveromyces lactis, genomic survey sequence.
ACCESSION AL425271.1 GI:12208465
VERSION AL425271.1
KEYWORDS GSS.
SOURCE Kluveromyces lactis.
ORGANISM Kluveromyces lactis.
REFERENCE 1 (bases 1 to 1021)
AUTHORS Souleir, J. L., Aigle, M., Artiguenave, F., Blandin, G.,
Bolotin-Fukuhara, M., Bon, E., Brottelet, P., Casaregola, S.,

```

```

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
COMMENT
FEATURES
source
misc_feature
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.:
Score:
Percent Similarity:
Best Local Similarity:
Query Match:
US-09-914-168-2 (1-919) x CNS071R5 (1-1021)
Oy 424 AsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGlu 443
Db 744 AATTGGCATCTGACCGCATTAAGAAAAAATGGCCAAAGAACAAAGAGAGGCGACAC 685
Oy 444 GlnSerSerSerArgTrpGluProAlaGlnValAspGluSerThrLeuGluProVal 463
Db 684 AGTAACAATAAATCAAAAGTCTGCTGCTCAGCAAAAGATGTCAAGGCTACAGAGCAGTC 625
Oy 464 IleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSer 483
Db 624 CAAGAACTGGTCTCTGAAATATGATGC----- 598
Oy 484 AlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyr 503

```

de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B., Malpertuy, A., Neuvéglise, C., Ozier-Kalogeropoulos, O., Potier, S., Saurin, M., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M., Winkler, P. and Weissbach, J.

Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies

2 (bases 1 to 1021)

11152876

20584721

11152876

11152876

2 (bases 1 to 1021)

Bolotin-Fukuhara, M., Toffano-Nioche, C., Artiguenave, F., Duchateau-Nguyen, G., Lemaire, M., Marmisise, R., Montrocher, R., Robert, C., Termler, M., Winkler, P. and Wesolowski-Louvel, M.

Genomic exploration of the hemiascomycetous yeasts: 11.

Kluveromyces lactis

FEBS Lett. 487 (1), 66-70 (2000)

20584721

11152866

3 (bases 1 to 1021)

Genoscope.

Direct Submission

Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 Evry cedex, FRANCE. (E-mail : segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

This GSS is part of a random genomic sequencing program of thirteen yeast species: *Saccharomyces bayanus* var. *uvarum*, *Saccharomyces exiguus*, *Saccharomyces servazii*, *Zygosaccharomyces rouxii*, *Saccharomyces kluyveri*, *Kluveromyces thermotolerans*, *Kluveromyces lactis* var. *lactis*, *Kluveromyces marxianus* var. *marxianus*, *Pichia angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*, *Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.

Location/Qualifiers

1..1021

/organism="Kluveromyces lactis"

/strain="CLIB 210"

/variety="lactis"

/db_xref="taxon:28985"

/clone="BA0AB004F03"

/clone.lib="BA0AB"

complement(1..>618)

/note="similar to *Saccharomyces cerevisiae* ORF YOR216c [RUD3 ; suppressor of us01-1 transport defect]"

/evidence="not-experimental"

Db 597 -----GCTAACGCTGTCCG----- 583
Qy 504 AsmepProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle 523
Db 582 ---GTACCGGATGAT-----TCTCTAGAAACTCTGATGGTGGTGCACCATGCACT 535
Qy 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspIleUser 543
Db 534 GTG-----GAAGCGGTGTCTACTGTGAAGAGATGACAGCATGCAAACT 487
Qy 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArg----- 559
Db 486 GAACCGAGGAAACAGCAATCATCACGACGACCAACAACTTAATGATGAGAGGACAGACCG 427
Qy 560 -----LysThrProAlaAspValTyrGln 567
Db 426 CACAGTCATTCTGAGGAGACAGCATCTACAGCGCTTTAAAGAGAAAGATAGTACGAA 367
Qy 568 SerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyLeuIle 587
Db 366 TTGAAGTATACACTCTTTGAGTAAATCTCATCTATGAGACCGTGTTCAGTCAATG 307
Qy 588 GlyLeuGlyTyrGlySerAspThrGlyThrArg---LeuValThrLysPheGluHisAsn 606
Db 306 AAGCAAGCTCAACAGCAAGTGGAAACACTCTGTGACAGTGTGAATGAGATGAGATCAG 247
Qy 607 LeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLys 626
Db 246 AATATTAATTTAAACAGACAGTAAATCTGCTGATTCAGAAATCCAAAGAGACAGCA 187
Qy 627 GlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAspGlnLeuArgAla 646
Db 186 ACATTTCTATT-----TTGATGATGAAATC---GTC 157
Qy 647 ThrLeuGlyTyrGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThr 666
Db 156 AATCTCAATTAAGATGTAAGTGAATTTGATGATGATGATGATGATGATGATGATGATG 97
Qy 667 ArgThrLeuGluHisGluIleSerArgSerIleGlnAsnGlyGlyTyrAsnArgThr 686
Db 96 AAAGATTTGGAATCTTAAGTTGACGAGCA-----TCTACGAGAGAGTCTTAACATTACT 43
Qy 687 TyrSerLeuArgTyrArgLeuAspLysLeuLysThrGln 699
Db 42 TCTTGTCTTAATCTGAAGTCAAAAGCTTAAATCTCAG 4
RESULT 25
AM342205 1364 bp mRNA linear EST 01-JAN-2001
LOCUS
DEFINITION GtHESTR91 Gulliardia theta Lambda Zap II cDNA library Gulliardia
ACCESSION AM342205
VERSION AM342205.1 GI:12000580
KEYWORDS EST.
SOURCE Gulliardia theta.
ORGANISM Gulliardia theta.
REFERENCE 1 (bases 1 to 1364)
AUTHORS Fruhnholz,M., Duebel,J., Wastl,J., Zauner,S. and Maier,U.-G.
TITLE EST Database of the cryptomonad alga: Gulliardia theta
JOURNAL Unpublished (2000)
COMMENT Contact: Maier, U.-G.
Department of Cell Biology and Applied Botany
Philipps-University Marburg
Karl-von-Frisch-Strasse, D-35043 Marburg, Germany
Tel: +49 6421 282 2057
Fax: +49 6421 282 1543
Email: maier@mailer.uni-marburg.de.
FEATURES
source 1..1364
location/Qualifiers
/organism="Gulliardia theta"
/strain="CCMP327"
/db_xref="taxon:55529"

/clone="Est25_39_46_8"
/clone_id="Gulliardia theta Lambda Zap II cDNA library"
/lab_host="SOLR"
/note="Vector: Lambda Zap II; Site_1: EcoRI"

BASE COUNT 383 a 286 c 343 g 348 t 4 others

ORIGIN

Alignment Scores:

Pred. No.:	1.1	Length:	1364
Score:	102.50	Matches:	71
Percent Similarity:	36.66%	Conservative:	43
Best Local Similarity:	22.83%	Mismatches:	98
Query Match:	2.17%	Indels:	99
DB:	10	Gaps:	15

US-09-914-168-2 (1-919) x AM342205 (1-1364)

Qy 261 GluValAspValIleIleHisAspLeuGlyLuproValTyrIleAspTyrArgAlaVal 280
Db 168 GACGTTGATACTATCATCATCAAGGATTTG-----GTAGACATTCCTGTAGTG 212
Qy 281 GluValArgGlyGluGlyAlaAspAsp---LysAlaPheThrThrValAlaAspGlu--- 298
Db 213 CTGACGCGGACACAGATGAGACAAATTTGAAGGAGAGAGTGGCGATCGCTGTGAAAC 272
Qy 299 -----ValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThrLysLys 316
Db 273 AGAATTTCTGATATTTCTTCTGGGATGTGGACAGTCAAGACCAACCAATCTTCCGT 332
Qy 317 AsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgThrLeuAspArg 336
Db 333 GCATTGTGCTGAGGCGCGCTTGAG-----GACCGC 365
Qy 337 SerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGly 356
Db 366 ACAATCGAATGATGATGCTGCTCAACCTGGCAATCGCAATCGCAAGATATCTGTGA 425
Qy 357 ThrGlnTyrArgPheAspGlyValValPhePheThrIleAspProLysThrAsnGlnLeu 376
Db 426 -----TTTCATTTGCGCATTTTGTCTTCAGATGAGCAATACATATGCGCAACAGACT 479
Qy 377 ThrThrAspProAspLysLeuProValLysArgGluLeuLeuGlnLeuLeuThrVal 396
Db 480 CAGACGAGAGAGATGAGATCTCTGAATGACAGACCATCTTGAA----- 524
Qy 397 AsnMetGlyGluAlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnLeuIleAla 416
Db 525 -----GACCTTGAAGCT 536
Qy 417 ThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGln 436
Db 537 NACAG-----ATGATGAATTCGGAGACATT-----GAGGCGCATCGATCAAG 581
Qy 437 Asn-----AspGlnVal-----SerPhe 442
Db 582 AATACAGAGAAATGAATGATTTGTTCATTGATGATGATGATGATGATGATGATGATGATG 641
Qy 443 GluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluPro 462
Db 642 ACCTACGGGGCGTGTGCTGTCTGTGAAGGTGTTCAGAGGGAT-----CTTCTTCT 692
Qy 463 ValIleGlyThrValGluLeu-----ThrAspGlyIleLeu 474
Db 693 TTGATGGAAGATGTTTCATCAGACCAACAGTATGGAACAGTGAAGACACACCATATTCTT 752
Qy 475 MetAspIleSer----- 478
Db 753 TTGCTGTCTGTGAGCATTTCTACGCGCAAACTTTCAGACTTGTCTCGTACGTCAA 812
Qy 479 -----ProIleGluThrSerAlaSerAlaSerAlaSerAlaSerAlaSerAlaSerAla 495
Db 813 GGAAGGCTGCCATTCGCGCTTGAATTTGAAGCTTTGACCAAGACAGACCTTTTCTTGATC 872

0y	496	AlaAlaValysAlaGlnHisLeuTyrAspMet-----ProAspArgVal	510
Db	873	CTTACCGTGGCCGGAGTGAATTTGATGATACACGACCAATTGCCCTCTGGAGACCGAATA	932
0y	511	LeuAlaLeuAsnHisAspAspGlyValAsnArg	521
Db	933	TGTACAGCTT-----GGTGTTCACAGA	953
RESULT 26			
LOCUS	BC019335	1867 bp	mRNA
DEFINITION	Homo sapiens, clone IMAGE:3607338,		linear
ACCESSION	BC019335		
VERSION	BC019335.1		
KEYWORDS	HTC.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 1867)		
JOURNAL	Strausberg, R.		
	Direct Submission		
	Submitted (13-DEC-2001) National Institutes of Health, Mammalian		
	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
REMARK	NIH-MGC Project URL: http://mgc.ncl.nih.gov		
COMMENT	Contact: MGC help desk		
	Email: cgabbs-remail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Rubin Laboratory		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Genome Sequence Centre,		
	BC Cancer Agency, Vancouver, BC, Canada		
	info@cgsc.bc.ca		
	Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,		
	Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,		
	Letlicia Hsiao, Martin Krzyzanski, Reta Kutsche, Oliver Lee, Soo		
	Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven		
	Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline		
	Scheil, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Stolt,		
	Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,		
	George Yang, Scott Zuyderduyn, Marco Marra.		
FEATURES	Clone distribution: MGC clone distribution information can be found		
	through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
	Series: IRAL Plate: 18 Row: F Column: 1		
	This clone was selected for full length sequencing because it		
	passed the following selection criteria: matched mRNA gi: 4758689		
	This clone has the following problem: frame shifted.		
	location/qualifiers		
SOURCE	1. 1867		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:3607338"		
	/tissue_type="uterus, endometrium adenocarcinoma"		
	/clone_lib="NIH MGC_44"		
	/lab_host="DH10B-R"		
	/note="Vector: pOTB7"		
BASE COUNT	554 a 465 c 556 g 292 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	1 92	Length:	1867
Score:	102.50	Matches:	134
Percent Similarity:	31.5%	Conservative:	90
Best Local Similarity:	18.90%	Mismatches:	249
Query Match:	2.17%	Indels:	237
DB:	11	Gaps:	32
0y	43	ProAlaHisAspThrAlaLeuAsnGlnAlaValysAlaGlnHisProProValLeuThr	62
		US-09-914-168-2 (1-919) x BC019335 (1-1867)	

Db	63	CCTTCCACCCCGACCCGACACACCGCGGCGCCACCTGCGGCCCGCCCGCCCTGTCGGCGCG	122
OY	63	ProGluGlnIleGlnAlaIaArgLeuAsnAlaIaGlyLeuAsnAlaIaProGlnSerGln	82
Db	123	CCCGAGCCACGCGCGCGAGCCCTCCCGCGCGCGCTCGCTCGGCCCGGAAAGCGCA	182
OY	83	AlaLeuAspValValAsnPheAspAspGlnSerProIleSerArgIleGlyGlnIleSer	102
Db	183	GCCTTCACTTAGCGCGCAGT-----GCTCCGCTCCGCGGACA---GAGCGCGCG	230
OY	103	ProProLeuGlyLeuAspMetSerValIleGluGlnThrThrProLeuSerLeuGluGln	122
Db	231	CCCCCT-----GGCCCGCGCCGCGAGGGCTCCCGCG--CGGCTCCCGAG	274
OY	123	LeuPheAlaGlnIleUserThrGluMetGlyIleAsnProAsnAspTryIleProGluTyr	142
Db	275	CATTTCGG-----CCGGGTGAGCGGGCCGAGCCCG	307
OY	143	GlnGlyGlnGlnProAsnSerGluValValValProProThrIleGlnGluProGlyLeuPro	162
Db	308	CAGATGACACGCGCCCGCGCGCG-----TCMAAGCCG	340
OY	163	GlyLeuIleIysArgLeuTyrAlaArgLeuPheAsnAspGly-ValAsnIysVal-----	180
Db	341	GGA-----GATGACTGTTTAGCCCGGAGACCGAGAGCGCGCA	382
OY	181	ProArgLeuLysAlaIysPheTyrGlnSerSerGlnSerGlnIleThrSerAlaIleGln	200
Db	383	AGCCCGCTCCCTCAAAAGCGGGCGCGCGCGGAGGCTCGGAGATCCGATGAAAGA	442
OY	200	YSerSerHisGlnIleSthIrrGluProTyrAlaAsnIleLysAlaIaLeuIleAsnIleTh	220
Db	443	GCTGACCGCGAGAGAGAGAGAGTAAAGACAGCACAGAA-----AAAGATTTTAC	493
OY	220	ArgGlnSerAlaMetAspLeuAsnGly-----SerIleProArgLeuArgGlnThrAlaIle	239
Db	494	TGAGAGAGGGCTGCTGTAAATCGCGGGCTCTGTCGAGCCAGCGCTGGCTCTCGGGTGG	553
OY	239	ValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIle-----	254
Db	554	GACTTCCCTCTCGGAGA---GGCAGCGGAGACCTCCATCCATCCATCCACACCGAGCATC	610
OY	255	-----IleArgAsnSerIleGlyGlnValAspValIleIleHisAspLeuGlyI	271
Db	611	CATCAGGGAATCAAGACCTCTTACGACAGAGTTGAAG-----	650
OY	271	uProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAspLys--	290
Db	651	-----AAATATAGAAGCGCATGTGTTCCAAATGCTCAGCTACATGAAAA	697
OY	291	-----AlaPheThrThrValAlaAspGluValProLeuIleuIleGlyAs	305
Db	698	GACAACTTCATGTACCGAGGTGATACCTTAAAGATATGTGGCGGAGCTTGAAGACA	757
OY	305	PValPheHis--HisGlyIleTyrGluThrLysAsnLeuIleGluAsnAlaSerAl	324
Db	758	GCTGGCTCAATCTTAGCGCGCATGACGAAGAAAAACAAGAAATTGTAAGGGAANAACA	817
OY	324	agLysIleGlyTyrPheAspGlyArgTyrPleuAspArgSerValAspValIleLeuProAs	344
Db	818	CGCCACACGATACTG-----	833
OY	344	PasnThrLysAspValSerLeuIleTyrAspThrGlyThrGluTyrArgAspGluVala	364
Db	834	-----CAATTTAGATTGCTGAGT	853
OY	364	I-----ValPhePheThrIle	369
Db	854	CAAGAGGCGCTCAAGACAAGAGGAAGATCGAGACAATGCAATTAATCTTAATTTC	913
OY	369	eAspProLysThrAsnGlnLeuThrIrrAsp-----Pr	380

LOCUS	BC018883	1876 bp	mRNA	linear	HTC 07-DEC-2001
Db	914	AGAAATAGCTACCAATGAGAGACTTCGACACACCTCAATATGTTGGATTACCAAG6TCC			973
Oy	380	CAAPlySLeuProValLysArgGluLeu-----LeuGluInLeuLeuThrValAsnMe			398
Db	974	TACCAAGATG---ACAAAAGAAAGACTTAAATGCTCCCTCAAGTCGACAGGAGTGGAGCCCT			1030
Oy	398	GLyGluLysArgLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrAr			418
Db	1031	AGGAAGACCCACTGAAGTGGAG-----GTGAAAATGAAATTCGTGGC-----			1073
Oy	418	gTyrPheAsnMetValAsnThrGluIleValPhe---ProGluArgGluInIleGlnAs			437
Db	1074	-----AATGGGGGAAAAGAAAGAAATCTGCCAATCTACTGAGAAAGAAACAACACACAGA			1126
Oy	437	nasplVal		SerPheG1	443
Db	1127	GGACACAGTGAAGAGCTGTGTGGACATAGAGTATTCCTGCTGCTGAGAAATACAGAGA			1186
Oy	443	uGlnSerSerSerSerArgThrGluPro-----AlaGlnValAs			456
Db	1187	CCAGAAATCTCTGTGAAGACACTGCCCATCTTCTTAGAACCTTAGCAGAGGCTTACTTAGTA			1246
Oy	456	pGluSerThrIleuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAs			476
Db	1247	GGAAACAGCTTCAAAACCAATCTTGAGAGCAGTTCCTCTCCCTGAA-----			1292
Oy	476	pIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAl			496
Db	1293	-----AACACAGTACA			1303
Oy	496	aAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValIleuAlaIle-----			513
Db	1304	GCTTGAATCAATGAGGCTCATGGGTGCACCAATATACAGACAGACACTCCCTTGAGCC			1363
Oy	514	-----AsnHisAspAspGlyValAsnArgSerI1			523
Db	1364	ATCCAACCTTGTCAGTGAATAGATGGTGGGAAACACACAGCAATGTGGGAGAGGACAC			1423
Oy	523	eLeuGlnArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSe			543
Db	1424	AGTGACTCAGGTGTAAGAGACAGCAGCAGACACTGGCTCTGCTTGAAGGCATATGTA			1483
Oy	543	rGluAsnGlu-----ValIleAspLeuProGluArgThr-----			554
Db	1484	TGACACACTTTATCATGATGACAAATGTAATGGTATGAGGGCCCCCAAGAGTTAGACACAG			1543
Oy	555	-----AlaLeuAlaAsnArgLysThrProAlaAspValItyrGlnSerLysLysVa			571
Db	1544	CACAGCGCATAGTTTAGAAGAAAGATTCACCAACAGGAGCAGTGAAGCCCAAGAGAGT			1603
Oy	571	pProleuTyrValIlePheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTr			591
Db	1604	TTCACCGGCAC-----AGTACAGAACTAGTGGGAT-----			1634
Oy	591	pGlySerAspThrGlyThrArgLeuValIThrLysPheGluHisAsnLeuIleAsnArgAs			611
Db	1635	-----CACAC-----GAGGAAGA			1648
Oy	611	pGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerLysLysGlyValLysLeuTy			631
Db	1649	GGGTGAAGAAACAGA-----TTAAG			1669
Oy	631	rAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrG1			651
Db	1670	GGAGCAAGAAACCAATTCACAGACAGAGTCTCTCTCCAGCAGCAACTGAGGCGCACTG			1729
Oy	651	nGlnGluValPheGlyHisSerThr			659
Db	1730	TCAGAGAGCGAGAGGTCCTCAAGTACA			1754
RESULT 27					
LOCUS	BC018883	1876 bp	mRNA	linear	HTC 07-DEC-2001

DEFINITION	Homo sapiens, clone IMAGE:3617322, mRNA.
ACCESSION	BC018883
VERSION	BC018883.1 GI:17403016
KEYWORDS	HTC.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS	1 (bases 1 to 1876)
TITLE	†
JOURNAL	Strausberg, R. Direct Submission Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgabbs-remail.nih.gov Tissue Procurement: ATCC/DC/DTF CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/ Contact: nisc.mgc@nhgri.nih.gov Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, F., Legaspi, R., Lim, M., Maduro, Q.L., Masello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantirpop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgenev, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.
FEATURES	<p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRRL Plate: 12 Row: n Column: 11 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5174486 This clone has the following problem: no 5' EST match.</p> <p>Location/Qualifiers</p> <p>1..1876</p> <p>/organism="Homo sapiens"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="IMAGE:3617322"</p> <p>/tissue_type="Skin, melanotic melanoma."</p> <p>/clone_lib="NIH-MGC_20"</p> <p>/lab_host="DH10B-R"</p> <p>/note="vector: pORF7"</p>
BASE COUNT	556 a 467 c 561 g 292 t
ORIGIN	
Alignment Scores:	
Pred. No.:	1..94 Length: 1876
Score:	102.50 Matches: 134
Percent Similarity:	31.59% Conservative: 90
Best Local Similarity:	18.90% Mismatches: 249
Query Match:	2..178 Indels: 237
DB:	11 Gaps: 32
US-09-914-168-2 (1-919) x BC018883 (1-1876)	
43	ProAlaHisAspHrAlaIleAsnGlnAlaLysAlaGlyAsnProProValIleuEuthr 62
72	CGTTCCACCCCGAGCCGAGCTACGCCCGCGGCCACCTCGGCCCCCTGCGGCCGC 131
63	ProGlnGlnIleGlnAlaArGLeuAsnAlaIaGlyLeuAsnAlaLysProGlnSerGln 82
132	CCCGAGCCCGAGCCCGAGCCCGCTCCCGCGCGGCTGCTCCTCGGCCCGAGACGCGA 191
83	AlaLeuAspValValAsnIleAsnAspAspGlnSerProIleSerArgIleGlyGlnSer 102
192	GGCTACCTTACGAGCGGCGAGT-----GGCTCTCTCCCGAGACA--GAGCGCGC 239

OY	103	ProLeuGluLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGlu	122
Db	240	CCCCCT-----GGCCGGGCCCCGGAAGGCTCCCGGC-6CGGTCCCCGAG	283
OY	123	LeuPheAlaGlnGluSerThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyr	142
Db	284	CATTTCGCG-----CCGGGTGAGCGGGCCGACCCGG	316
OY	143	GlnGlyGluGlnProAsnSerGluValValProThrThrLeuGluProGluLysPro	162
Db	317	CAGATGACGCCGCCGCCGCCGCC-----TCAAAAGCCG	349
OY	163	GlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGly-ValAsnLysVal-----	180
Db	350	GGA-----GATCGACTGTGGACCCCGGAGCGCAGACAGCTGGCGGA	391
OY	181	-ProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyThrSerAlaIleGlu	200
Db	392	AGCCGGCTGGCTCGCAAAAGCGGGGGCGCGCGGAGGCTCGAGATCCGATGACGA	451
OY	200	ySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeuGluAspIleTh	220
Db	452	GCTGAGCGCCGACGAGAGGAGGAGGTAGACAGACACCCAGAA-----AAAGATTTCAC	502
OY	220	rGlnGluSerAlaMetAspLeuAsnGly--SerIleProArgLeuArgGlnThrAlaLe	239
Db	503	TGACAGAGCGCTCTGTACACATGCCGGGCTGTGTGACGCCACGCTGGCTCTCTGGTGG	562
OY	239	ValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIle-----	254
Db	563	GACTTCCTCTGGAGA---GGACGGAGACACCTCATCTCATCGACACGAGCGATC	619
OY	255	-----IleArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGlu	271
Db	620	CATCAGGAAATCAAGGACTCTAGCAGAAAGTTGAAAG-----	659
OY	271	uProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAspLys-	290
Db	660	-----AAATATAGAAGGCTATGGTTTCCATGCTGCTACGTACAAATGAAGAA	706
OY	291	-----AlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAs	305
Db	707	GACAAACTCATGTACACAGGTGTATACCCCTAAAGATATGTTGGCCGAGACTTGAAGACA	766
OY	305	pValIlePheHis--HisGlyLysTyrGlnThrLysLysAsnLeuIleGluAsnAlaSerAl	324
Db	767	GCTGGCTGCAATCTTAGCGCGGACGTACGAGAGAGAAACAAAGAATTTGAAAGGGA	826
OY	324	AGLHisGlyTyrPheAspClyArgTrpLeuAspArgSerValAspValIleLeuProAs	344
Db	827	CCCCACACGTATACTG-----	842
OY	344	pAsnThrAlaAspValSerLeuIleTyrTrpThrGlnTyrGlnTyrArgPheAspGluVa	364
Db	843	-----CAATTCACATTCGTGGAAGT	862
OY	364	I-----ValPhePheThrII	369
Db	863	CAGAGAGCCCTGAGCAAGAGAGAGAAATGCTCGAGAAACATGGAATATATCTTAAATTC	922
OY	369	eAspProLysThrAsnGlnLeuThrThrAsp-----Pr	380
Db	923	AGAAATAGCTACCAATGAGACAGACTTCCGACACCCATCATATATGTTGGATACCAAGATCC	982
OY	380	oAspLysLeuProValLysArgGluLeu-----LeuGlnGlnLeuLeuThrValAsnMe	398
Db	983	TACCAACATG---ACAAAGAGAGATTAAATGCCCCCTCAAGTCGACAGCGGATGGGACCCCT	1039
OY	398	tGlyGlyAlaIleTyrTrpAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThrAr	418
Db	1040	AGGAAGACCCAGTGAAGTTGGAG-----GTGAAAAATGAATGTGTGGCG-----	1082

QY	418	gtyrphesnmctvlaasnthglulevalphe---	ProgluIarvggluInleGlmas	437
Db	1083	-----AATGTGGGGAAGAAGAGAAATTTCTGCACATATCTAGAAAGAACACACACAGA	1135	
QY	437	nasplInval-----	-----SerpheG1	443
Db	1136	GGACACAGTGAAGACACTGTGTGCACATAGACATATTCCCTGCTGTGTAGAAATPACGAGAGA	1195	
QY	443	uGInserSerSerSerATgThnGlupro-----	-----AlaGlnValas	456
Db	1196	CCAGAAATCTCTGTAGACACACTGGCCCCATTCCTTAGGAACCTTAGCAGAGGTGCTTACCTATGA	1235	
QY	456	pGluserThreugluProValIleGlThrValGluLeuthraspglyIleleumetas	476	
Db	1256	GSAAACAGGTCAAAACCCAAATTTCTTGAGAGCAGGTCTCTCCCTGTAA-----	1301	
QY	476	pLieserProIleGlupheserAlaSerAsnLeuIleGlnasplysLeuasnLeuValAl	496	
Db	1302	-----	-----AACACAGTACA	1312
QY	496	aAlaLysAlaArgHisLeuTyTrAspMetProaspsparValleuAlaIle-----	513	
Db	1313	GSTTAGTCAAAATGAGTGCATGGGTGCACCAAGATGACAGCAGCAGACTCCCTTTGAGCC	1372	
QY	514	-----	-----AsnHisaspspglyValaAsnArgSerI1	523
Db	1373	ATCCAACTGTGGAGTGACTTAGATGGTGGGACACACAGAGATGTGGGAGAGCAGC	1432	
QY	523	eLeuIdIyrrgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSe	543	
Db	1433	AGTGCCTCAGGTTGAGAGCAGCAGCCACAGTGGCTCTGTGTCTTTAGGCGCATAGTA	1492	
QY	543	rgLusnsglu-----	-----ValIleaspleuProGluIarGthr-----	554
Db	1493	TGACACAGTATTATCATGTATGCACAAATGTATGGTAGAGGTCCCCCAAGATTAGAGACAG	1552	
QY	555	-----	-----AlaLeuAlaAsnArgLyThrProAlaAspValTyGlnSerLySlyVa	571
Db	1553	CACAGGGCATATTTAGAAGAAATTCCACCACAGCAGAGCAGCTGAGCCCAAGAGAGT	1612	
QY	571	lProLeuTyrrValPheValAlaSerAspLySproArgAspGlyInleGlyLeuGlyTr	591	
Db	1613	TCCACCGCAC-----	-----AGTACGAAAGTGTATGGAT-----	1643
QY	591	pGlySerAspThrglyThrArgLeuValThrLySpheGluHisasnLeuIleasnArgAs	611	
Db	1644	-----	-----CACAA-----GAAGAAGA	1657
QY	611	pGlyTyrrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLySglyValTySLeuTy	631	
Db	1658	GGGTACAGAACACAGA-----	-----TTAAG	1678
QY	631	rAlaThrLySProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrrG1	651	
Db	1679	GGACAGAAACCAATCAACAGACAGAGATTCTCGTTCTCCACAGAGAACTGAGCGCAACTG	1738	
QY	651	ngIngluValPheglyHisSerThr	659	
Db	1739	TCAGGAGCGACAGGTCCCAAGTACA	1763	
RESULT	28			
LOCUS	BI086482/c			
DEFINITION	60284975221 NIH_MGC_10 Homo sapiens	1014 bp	RNA	linear EST 20-JUN-2001
ACCESSION	BI086482			
VERSION	BI086482.1			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 1014)			

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgap@b-remail.nih.gov
 Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: Incyte Genomics, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>
 Plate: LAM11008 row: h column: 14
 High quality sequence stop: 463.

FEATURES
 source location/Qualifiers
 1..1014

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="491293"
 /clone_lib="NIH-MGC_10"
 /cell_line="MGC36"
 /lab_host="DH10B"
 /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.5 kb. Library prepared by Life Technologies."

BASE COUNT 213 a 325 c 299 g 177 t
 ORIGIN

Alignment Scores:

Pred. No.: 0.749 Length: 1014
 Score: 102.00 Matches: 60
 Percent Similarity: 30.77% Conservative: 20
 Best Local Similarity: 23.08% Mismatches: 87
 Query Match: 2.16% Indels: 93
 DB: 13 Gaps: 11

US-09-914-168-2 (1-919) x B1086482 (1-1014)

QY 573 LeuTyValPheValAlaSerAspLysProArgAsp-----GlyGln 586

DB 603 CTCGGCGTCTTCTGAATAGACCGGTTCCCGGACAGAGGCGCAATGACCCGGTCCCGAG 544

QY 587 TleGlyLeuGlyTTPcLysSerAspThrGlyThrArgLeuValThrLysPheGlnHisAsn 606

DB 543 GTCAGAGCTTGGGTG-----529

QY 607 LeuIleAsnArgAspGlyTyrglnAlaGlyAlaGluLeuArgLeuSerGlnAspLys 626

DB 528 -----CCCCAAGGTTTTCAGATTGCTGTCAGAGCCAGA-----496

QY 627 GlyValLysLeuTyAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAla 646

DB 495 -----AACCTGGTTCTCCGCTGT 478

QY 647 ThrLeuGlyTyrglnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThr 666

DB 477 CTTTGGGGGTCGCCGA-----GGGCATTCTACGTGGGAGAGAGGAATGGAAGTT 427

QY 667 ArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGlyLysTrpAsnArgThr 686

DB 426 -----TGGACCAAGTAA 415

QY 687 TyrSerLeuArgTyArgLeuAspLysLeuLysThrGlnAlaProGluThrTrp---705

DB 414 CTCCTCTGTCAGCCCGCTGCAAGCTTGGCCCTTCAGAGGACGCTCGGTGACTTGGTTC 355

QY 706 -----GlnAspLeuPro-ValAspPheValAsnGlyLys 716

DB 354 AGCCCTCCCGTCGATCCAGAAATGCGGATGGTCCGGGACCACTTCATATGCTGAAT 255

QY 716 sPseRgGlnGluAlaLeuLeuAlaGly-----ValAlaValHisLysThrTyrA 732

DB 294 GCCGGCCCAACCTGAGCTCTCGCAGGCTGAAGACCCGGCGGCCGCCACCTTCGTGT 235

QY 732 lAlaAspAsnLeuValAsnProMetArgGly-----TyrArgGlnArgTyrSerLeu 750

DB 234 GGT-----CCGACCGTGGGACGCGCACGATGGCCGATGGACCCG 190

QY 750 uValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIle 770

DB 189 AGCGGGGACAGCGGGCTGATTGCG-----GGCGCCCTTGGCTTGGC 148

QY 770 rGlyValTyrSerPheGlyAspAsnAlaTyrglySerAsnArgAlaHisGlnMetThrG 790

DB 147 GGGCCATTACGTTGCGGA-----TCTTACAGGCGGACGAGCATGTGAACCACTGTGC 94

QY 790 yGlyIleGlnAlaGlyTyrIleThrSerAspAsnHisValProTyArgLeu 809

DB 93 CACGCCCGCTGCCATCTTGTGGAAGTGGGCTTCAAGACCATGCAATCGGCTG 36

RESULT 29

AK013402 2039 bp mRNA linear HTC 19-JAN-2002
 LOCUS AK013402
 DEFINITION Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length

enriched library, clone:2810473H05:homolog to
 PHOSPHOCETYLTGUCOSAMINE MOTASE (PC 5.4.2.3) (PAGM)
 (ACETYLTGUCOSAMINE PHOSPHOMUTASE) (N-ACETYLTGUCOSAMINE-PROSPHATE
 MUTASE), full insert sequence.

ACCESSION AK013402
 VERSION AK013402.1 GI:12850738

KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (strain: C57BL/6J) 10, 11 days embryo cDNA to mRNA,
 clone_lib:RIKEN full-length enriched mouse cDNA library

ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P., and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)

2 Carninci, P., Itoh, M., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL MEDLINE
 PUBMED 2049374
 11042159

TITLE

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multichannel sequencer
 Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL MEDLINE

PUBMED

REFERENCE

AUTHORS

OY	309	HisGlyLysTyrGluThrLysLysAsnLeu-----IleGluAsnAlaSerLaeGluHis	326
Db	717	TCAGTTAAAGTTGATGCATGCCAACGGCATAGGGGCGCTTGAAACCTAAGACAAATGGAACAC	776
OY	327	GlyTyrPheAspGlyArgTrpLeuAspArgSerValAspValIleLeuProAspAsnThr	346
Db	777	---TACTTC-----TCCGGGGCGCTGTGGTT-----	800
OY	347	AlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPhe-----	361
Db	801	-----CTGCTGTTTAATGATGGAGGCCAACGGCGCTCAATCACCTGTGGCT	848
OY	361	-----	361
Db	849	GCTGACTTTGTCAAAAGTACACAGAAACCCACAGGCAATTGAATGAAGTCCGGTGAG	908
OY	362	-----AspGluValIlePhePheThrIleAspProLys	372
Db	909	ACATGCTGCTCTTCGTGGATGGAGATGGGACAGAGATCGTGTATTACTACTGTATGCAGAT	968
OY	373	ThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeuGln	392
Db	969	GCTCACTTTCATCTCATATGATGAGACACAAAGATAGCGACCTTAATTAAGCATTTCTTAA	1028
OY	393	LeuLeuThrValAsnMetGlyGluValTyrAsnLeuGlnAlaValArgLalaLeuSerAsn	412
Db	1029	GAGCTACTCTTGTGAGATTGGAAGAAAGTGTGAACCTCGGAGCTGTACAGACACATATGCA	1088
OY	413	AspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArg	432
Db	1089	AATGGAAGTTCACGCGGAGTACCTGTGAAGAAGTTATGAAGGTACCTGTGTATTGCAACAA	1148
OY	433	GluGlnIleGlnAsnAspGln-----ValSerPheGlu	443
Db	1149	ACTGGTGTAAACATTGTGCATCAACAAGGCTCAAGATTGACATCGAGATTATTATTGA	1208
OY	444	GlnSerSerSerArgThrGlnProAlaGlnValAspGlnSerThrLeuGlnProVal	463
Db	1209	GCGAACGGGCATGGAACACAGCACTGTTCACTGAACAGTTGAAGTGAACATAAAGACACTA	1268
OY	464	IleGluThrValGluLeuThrAspGly-----IleLeuMetAsp	476
Db	1269	GCCCA-----GAACCTAGACGATGAAGAAAGAAAGACCGACGAGCGCTGCGAGCATC	1322
OY	477	IleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAla	496
Db	1333	ATCGACTTGTTCACACGAGCACTCGAGATGCTATTCTGCATGTGTGGTATTTGAGCT	1382
OY	497	AlaLysAlaArgHisLeuLysArgMetProAspArgValLeuAlaIleAsnHisAsp	516
Db	1383	ATCCGAGGCTCTG-----AAGGGCTGACTCTATACAGCANCTGG	1418
OY	517	AspGlyVal-----AsnArgSerIleLeuGlyArgIleSerAsp-----	529
Db	1419	GATGCTATTATGTCGATCTTCTTAACAGACAACTAAAGTTAAAGTTCGCGAGGAGACA	1478
OY	530	-----AlaValSerLalaValAlaArgAlaIleLeuProAspGlnSerLysGln	546
Db	1479	GTTATTAGCACACAGGATGCTAGAGACAAAGAGTCACTCCACAGACTCCAAAGAGCA	1538
OY	547	ValIleAspLeuProGluIleThrIleAlaLeuAla	557
Db	1539	ATCATGACCTGCTGAAGAAATACACACTTGGC	1571
RESULT 30			
LOCUS	A2927049	513 bp	DNA
DEFINITION	476.dist16c05.sl Saccharomyces castellii NR1_Y12630 Saccharomyces		
ACCESSION	A2927049		
VERSION	A2927049.1	GI:13497952	
KEYWORDS	GSS.		

ORIGIN	BASE COUNT	181 a	102 c	90 g	140 t
US-09-914-168-2 (1-919) x A2927049 (1-513)					
QY 372 LysThrAsnGluLeuThrThrAspProAspLysLeuProValLysArgLysLeuGlu 391	:	:	:	:	:
DB 5 AAAAATATATAAACACCACCAAAACCAAGT-----GAACTCTTGA 46					
QY 392 GlnLeuLeuThrVal---AsnMetGlyGlnAlaTyrAsnLeuGlnAlaValArgAlaLeu 410	: :	: :	: :	: :	: :
DB 47 AAGCTAGCTACTATTATGAAACAACGGT-----AAAGCTCTTA 82					
QY 411 SerAsnAsp-----LeuIleAlaThrArgTyrPheAsnMetValAspThrGluIle 427	: :	: :	: :	: :	: :
DB 83 AGGATGTGATGATTGATTCACAGGCCACGACTGCTTATGAGATTCGCCACC----- 136					
QY 428 ValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSer 447	:::: :	:::: :	:::: :	:::: :	:::: :
DB 137 -----CCCTGACACATCGAAGATGAAGAAATTCGGAGGATCGTCCACCAGCGTTGAAA 190					
QY 448 SerArgThrGluProAlaGlnValAspGlu-----SerThrLeuGluPro 462	: :	: :	: :	: :	: :
DB 191 TCGAATACCAAGATCTTAATGTGAGATCAAGATCTATATCCATCGTTACTCGCG 250					
QY 463 ValIleGluThrValGlnLeuThrAspGlyIleLeuMet-----AspIleSer 478	::: : :	::: : :	::: : :	::: : :	::: : :
DB 251 ATTATACAAAGGTTTCATCCGAAAAACAATTATATGTTGGTGATACGAATCAGACT 310					
QY 479 ProIleGluPheSerAlaSerAsn-----LeuIleGlnAspLysLeuAsnLeu 494					
DB 311 ATTATGGAATTTAGGGGATTAATTCAGAATATTAAGTGCCTTGATAGACCTTCACCTG 370					
QY 495 ValAlaAlaLysAlaArgHisLeuTyrAsp-----MetProAsp----- 507	::: : :	::: : :	::: : :	::: : :	::: : :
DB 371 ACCAATTCACGACTCTCCACTTTATGATTAACCTTCGATGCACACCGAAATATTAAT 430					
QY 508 -----AspArgValLeuAlaIleAsnHisAspAspGly 518	: :	: :	: :	: :	: :

```

DB      431   TC0CGCACACGAATTCGTCA---AACCATCATTGAGGA    466
RESULT  31
LOCUS   BC028159                                     2340 bp     mRNA       linear      HTC 01-MAY-2002
DEFINITION Homo sapiens, similar to KIAA0076 gene product, clone IMAGE:5240443, MRNA.
ACCESSION BC028159
VERSION   BC028159.1
KEYWORDS GI:20380216
SOURCE   HTC.
ORGANISM Homo sapiens.
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2340)
AUTHORS Strausberg,R.
TITLE    Direct Submission
JOURNAL Submitted (08-APR-2002) National Institutes of Health, Mamalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK   NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT   Contact: MGC help desk
          Email: cgabbs@remail.nih.gov
          Tissue Procurement: Life Technologies, Inc.
          cDNA Library Preparation: Life Technologies, Inc.
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
          DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
          Gaithersburg, Maryland.
          Web site: http://www.nisc.nih.gov/
          Contact: nisc-mgc@nhgri.nih.gov
          Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Bikesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-T., Karlins,E., Latic,P., Legaspi,R., Maduro,Q.L., Mastello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantirlop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,L.H., Walker,M.A., Wetherby,K.D., Wiggin,L., Young,A., Zhang,J.-H. and Green,E.D.
FEATURES             Source
         source              1..2340
                               /organism="Homo sapiens"
                               /db_xref="taxon:9606"
                               /clone="IMAGE:5240443"
                               /tissue_type="Brain, fetal, whole pooled"
                               /clone_lib="NIH_MGC_121"
                               /lab_host="DH10B"
                               /note="Vector: pCMV-Sport6"
BASE COUNT            519 a        631 c        721 g        469 t
ORIGIN
Alignment Scores:
Pred. No.:           4.35               Length:                2340
Score:               101.00             Matches:                 109
Percent Similarity:  32.908             Conservative:           70
Best Local Similarity: 20.04%           Mismatches:            199
Query Match:         2.14%              Indels:                 166
DB:                  11                   Gaps:                    27
US-09-914-168-2 (1-919) x BC028159 (1-2340)
QY      84   LeuAspValIvalAsnPheAspAspGlnSerProIIeserArgIlleglyClngInserPro 103
          ||||| :|||||
          |11111 :|||||
          |-----AGCATTCTGGAGCTGGAGCTGGAGAGCCAAAGCTGCC 908

```

QY	104	Pro-----	-----	LeuGlyLeuAspMetSerVal	111
Db	909	CCATGTCGCACAGAGAGAGAAAA	CCGGGGACAGCGGGAACCTGGAACTTCAGCATGGCTGTG		968
QY	112	IleGluIuThrThrProLeuSerLeuGluGluLeuPheAlaGlnGluSerThrGlnMet			131
Db	969	GGCAACCTCATCTCTGAGCTT	GGCGGAGACATGGGCTGGAACTCAGCGAAAG		1028
QY	132	GlyIleAsnProAsnAspTryIleProGluTryGlnGluGlnProAsnSerGluVal			151
Db	1029	GGCATGTCACCT-----	CCCGCGCAACCGGCTGCATC		1061
QY	152	ValValPro-----	ProThrLeu-----	GluProGlu	160
Db	1062	TTTCAGCGCTACATTTTCAGGCCCC	ACGCTTTACTCCACCATTGTTCACACCCCGACA		1121
QY	161	LysProGlyLeuIleuLeuAspArgLeuTryAlaArgLeuPheAsnAspGlyValAsnLysVal			180
Db	1122	AGACAAAGGTGGGTC-----			1136
QY	181	ProArgLeuLysAlaLysPheTryGlnSerSerGlnSerGlyGluThrSerAlaIleGly			200
Db	1137	-----	TTCCGCCAGCGCTCGAATTTCCAGCCGTAGTGCGCTATGGA		1178
QY	201	SerSerHisGlnLysThr-----	GluProTryAlaAsnLysAlaLysAlaLysAlaLysAlaPhe		219
Db	1179	GAATATGTGCAGACACATCGACAGCCAGGAGATGGCGATGCGG-----	ATGCTGATGATATAT		1235
QY	220	ThrGlnGluSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeu			239
Db	1236	GAGAGATCATCGCT-----	GGGACAGAGGCGACTCCGGCAGAC-----		1277
QY	240	ValAlaAlaArgAlaValGlyTryTryAspIleAspLeuSerIleIleArgAsnSerIle			259
Db	1278	-----	AAACAGCGCATT		1289
QY	260	GlyGluValAspValIleIleHisAspLeuGlnGluProValTryIleAspTryArgAla			279
Db	1290	CCCCCTGGCGAGGTCTTTCGACATCGACAGCGCGCATTTACTGGGTGCACCTGGGCACATG			1349
QY	280	ValGluValArgGly-----	GluGlyAlaAspAspLysAlaPheThrThrValAlaAsp		297
Db	1350	CTGCAATGCTCTGGGCGCTCGACGAACCCAGTCAAGATAAAGCTTCACAGAGCTGTGGACAG			1409
QY	298	GluVal--ProLeuLeuIleGlyAspValPheHisHisGlyLysTryGluThrLysLys			316
Db	1410	GGCGCAGGCGCTACTCTGTGGGCGACAGCATTT-----			1442
QY	317	AsnLeuIleGluAsnAlaSerAlaGlnHisGlyTryPheAspGlyArgTyrLeuAspArg			336
Db	1443	-----	CCCTCTGGGAGTGGAACTCATATGGATGGGCTGAC-----		1478
QY	337	SerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTryAspThrGly			356
Db	1479	CGTTTGCGGTACCTCCAGCCCGAAGCTCAGAAAGATGACAGAGTGGAGATAT-----	CTG		1532
QY	357	ThrGlnTryArgPheAspGluValValPhePhe-----	ThrIleAsp-----		370
Db	1533	ACCACGGCTGAATGGTGGAGCTGCTTTTCTTATCAAAAAGTTGACCTGTGTGCAGAG			1592
QY	371	--ProLysThrAsnGlnLeuThrThrAspProAsp-----			381
Db	1593	CAGCCAAATTTTCCGAATCTTTTGGAGAACCTGGATGACACCCTGGGTGMAAAGCCCTA			1652
QY	382	--LysLeuProValLysArgGluLeuGlnLeuGlnLeuLeuThrValAsnMetGlyGlu			400
Db	1653	GGTGAATCTCTGTGTCGCGGAAATGGCCGACAGTCTGTGAGAGTT-----	CTCAGTAGT		1709
QY	401	AlaTryAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTryPhe			420
Db	1710	CGATTGTGGGGCAGACACTGTCATATGATCTGCTCACTCCAGATCTTACACCAAGATATGGG			1769

Y	421	AsnMetValAsn	-----	ThgGluIleValPheProGlu	431
Db	1770	CTGCTGCTATGAACAACAGCAGCTGCTGCTACTTCACAGAAATCACTCTGATCCCAAGAT			1829
Y	432	ArgGluGlnIleGlnAsnAspGlnValSerPhe	---GluGlnSerSerSerArgPhe	450	
Db	1830	CCAGAAAGAGAGTCCAAATCGAGAGCCAGCTTCTCAGAGAAAGACTGAGTCCCTCAA			1889
Y	451	GluProAlaGlnValAspGluSerThrLeuGlnProValIleGluThrValGlu	-----	468	
Db	1890	GCAAAAGCCGAGGCGCTTAAGACAGAGGCGGACCCACCAAGACAAAGACGAGACCC		1949	
Y	469	-----	LeuThrAspGlyIleLeuMet	475	
Db	1950	ATGCAACAGAGTGAATTCACAGCTGTTTAACCAAGCTTCTGGTGATCAGAGGGAGTGACCTG		2009	
Y	476	AspIleSerProIleGluPhe	---SerAlaSerAsnLeuIleGlnAspIleAsn	493	
Db	2010	-----CCACATGAGATGAGAGAGCAGCCAGCTGAAATGGCCAGA	-----	2048	
Y	494	LeuValAlaValAlaLysAlaArgHisLeuThrAspMetProAspAspArgValLeuAlaIle	513		
Db	2049	-----GCCTGGGGGGTCCGCTCCGACAGCTCCCTGATCAGACATGTCGACACGGTC	2102		
Y	514	-----	AsnHisAspAspGlyValAlaAsnArgSerIleLeuGlyArg	526	
Db	2103	GTGGCCACTGTGCAGATATCCAGCTTGGACACAAACCTGCAGCTTTCAGGGCTCTCTCC	2162		
Y	527	IleSerAspAlaValSerAlaValAlaLys	-----AlaIleLeuProAspGlu	542	
Db	2163	CTCTTCAGAGCTGTGTGAGAGAGTCACTAGAGCGGAGCACCCCTGTGCTGCTCTACAGA	2222		
Y	543	SerGluAsnGlu	546		
Db	2223	TCGCTGACAGAGAG	2234		
RESULT 32					
AK011769					
LOCUS	AK011769	2936 bp	mRNA	linear	HTC 19-JAN-2002
DEFINITION	Mus musculus 10 days embryo whole body cDNA, RIKEN full-length enriched library, clone:261004J33:transformed mouse 373 cell double minute 1, full insert sequence.				
ACCESSION	AK011769	GI:12848103			
VERSION	AK011769.1				
KEYWORDS	HTC; CAP trapper;				
SOURCE	Mus musculus (strain:C57BL/6J) 10 days embryo cDNA to mRNA, clone:11b:RIKEN full-length enriched mouse cDNA library				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol.	303, 19-44 (1999)			
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE					
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res.	10 (10), 1617-1630 (2000)			
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE					
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsuoka, Y., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsushita, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.				

TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer

JOURNAL
Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE
20530913

PUBMED
11076861

REFERENCE

AUTHORS
4
Kawai, J., Shingawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Atakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishik, K., Kiyosawa, H., Kondo, S., Yamataka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kusuwada, T., Saito, R.,
Kodota, K., Matsuda, H., Ashburner, M., Batilov, S., Casavant, T.,
Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L.M., Staib, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, D., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Momtaz, P.,
Rang, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyo-oka, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kotsuki, S.,
and Hayashizaki, Y.

TITLE
Functional annotation of a full-length mouse cDNA collection

JOURNAL
Nature 409 (6821), 685-690 (2001)

MEDLINE
21085660

PUBMED
11217851

REFERENCE

AUTHORS
5 (bases 1 to 2936)
Aach, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,
Atakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hasegawa, T.,
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Horii, F.,
Hume, D., Imtani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Konno, H., Konda, M., Koyama, S.,
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishik, K., Nomura, K.,
Nito, A.H., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,
Schriml, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T.,
Sogabe, Y., Suzuki, H., Tagami, Y., Tagawa, A., Takahashi, F.,
Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamataka, I.,
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and
Hayashizaki, Y.

TITLE
Direct Submission

JOURNAL
Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.

FEATURES
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
15' GAGACGAGCAAGATCCAGACCTCTTTTCTTTTCTTTTCTTTT 3', cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 7.5 and subtraction to
Rot = 37.5. Second strand cDNA was prepared with the primer adapter
of sequence 15'
GAGACGAGCATTCGAGCTTAATTAATTAATTCCTCCCCCCCC 3'. cDNA was cleaved
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

LOCATION/Qualifiers
1..2936
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM.DB:2610044J23"

Db	Accession	Source	Organism	Reference	Authors	Journal	Comment
Db	679	CTGGGAGGCCGACGAACTGCATTAAGCTGGAGCGCCAGAGTACGCTCGMGGGGGCGAG	738				
Oy	323	-----AlaGluHisGlyTyrPheAspCylValArg-----	332				
Db	729	TGCCATATCTACGATATACCTCATTCGACAGACGAGCAGCTTGACAGAGTGCCTTCGAGCT	798				
Oy	333	-----TrpLeuAspArgSer-----ValAspValIleLeuProAspAs	345				
Db	799	CCGACATCCAGCTGGCGCCACACATCAGGACGACACCATTGCGCCATTCCTCGTGGCCAA	858				
Oy	345	nThrAlaAspValSer 350					
Db	859	CAAGCGACACTTGGCC 874					
RESULT 34							
LOCUS	Bf256095/c						
DEFINITION							
ACCESSION	Bf256095	800 bp	mRNA	linear	EST 22-OCT-2001		
VERSION	Bf256095						
KEYWORDS							
SOURCE	Bf256095.2	GI:13118036					
ORGANISM							
REFERENCE							
AUTHORS							
JOURNAL							
COMMENT							

DB	Accession	Length	Matches	Conservative	Mismatches	Indels	Gaps
US-09-914-168-2	(1-919) x BF256095 (1-800)	800	100.00	33.22%	140	65	9
QY	14 ProValAlaLeuAlaAlaTyrLeuProLeuMetThr-SerGlnAlaLeuAlaGlnGlnAs	33	1111111111	1111111111	1111111111	1111111111	1111111111
Db	789 CCGGGCGCACACCGCGCATCACCAAAACACACCCACCGCCACACGACCTGCGAGCA	730	1111111111	1111111111	1111111111	1111111111	1111111111
QY	33 nAsnProAlaAsnIleIleAsnHsValProAlaHsAspThrAlaIleAsnGlnAlaIu	53	1111111111	1111111111	1111111111	1111111111	1111111111
Db	729 CAATCCCGCC-----	720	1111111111	1111111111	1111111111	1111111111	1111111111
QY	53 sAlaGlyAsnProProValLeuLeuThrProGlnGlnIleGlnAlaArgLeuAsnAlaI	73	1111111111	1111111111	1111111111	1111111111	1111111111
Db	719 -----CCCCAAATAGTACCCCGCCCGCACAA-----ACCGGCAACTACAGCTGC	673	1111111111	1111111111	1111111111	1111111111	1111111111
QY	73 aGlyLeuAsnAlaIuProGlnSerGlnAlaLeuAspValValAsnHsAspGlnSe	93	1111111111	1111111111	1111111111	1111111111	1111111111
Db	672 CGGTCTCACCGCCCGACATCCCGCCATCTC-----GAGTC	634	1111111111	1111111111	1111111111	1111111111	1111111111
QY	93 rProIleSerArgIleGlyGlnGlnSerProProLeuGlyLeuAspMetSerValIleG	113	1111111111	1111111111	1111111111	1111111111	1111111111
Db	633 CCGCGCTCTCGACAGCAGCAGGTACCCCGCTCCCGCCAGCCAGCAACAGCATCCCGC	574	1111111111	1111111111	1111111111	1111111111	1111111111
QY	113 uGlnThrThrProLeuSerLeuGlnGlnLeuPheAlaGlnGlnSerThrGlnMetGlyI	133	1111111111	1111111111	1111111111	1111111111	1111111111
Db	573 ACCCACCGCCCGCTCTC-----	555	1111111111	1111111111	1111111111	1111111111	1111111111
QY	133 eAsnProAsnAspTyrIleProGluTyrGlnGlyGlnIleProAsnSerGluValAla	153	1111111111	1111111111	1111111111	1111111111	1111111111
Db	554 -CGCCCGCGCGCCATCCCGCCACACAGCGCCCGCCACCGCCCGCCAGCCAGCG	496	1111111111	1111111111	1111111111	1111111111	1111111111
QY	153 lProProThrLeuGluProGlnGlnProGlyLeuIleLeuArgLeuTyrAlaArgLeuPh	173	1111111111	1111111111	1111111111	1111111111	1111111111
Db	495 CCGCGCTCTCGCGCGCCCGCTCGCGCCCGCCCGCTCCCGCCATGGCCACCGACGCCAC	436	1111111111	1111111111	1111111111	1111111111	1111111111
QY	173 e-AsnAspIleValAlaAsnValProAlaGlyLeuAlaI-LysPheTyrGlnSerSerGln	192	1111111111	1111111111	1111111111	1111111111	1111111111
Db	435 CCGACCGCCCGACCGCGCGAGTCCGAGTCCCGCGAGCGCGTCCCATGATCCCGCCCG	376	1111111111	1111111111	1111111111	1111111111	1111111111
QY	193 SerGlyGluThrSerAlaIleGlySerSerHisGlnIuThrGluProTyrAlaAsnIle	212	1111111111	1111111111	1111111111	1111111111	1111111111
Db	375 CATGTCGCGGCGCCAGCGCTCGGCGCATCGCCGACGACGACACCA-----GCCGGTGG	319	1111111111	1111111111	1111111111	1111111111	1111111111
QY	213 LysAlaAlaLeuGlnAspIleThrGlnGlnIuSerAlaMetAspLeuAsnGlySerIlePro	232	1111111111	1111111111	1111111111	1111111111	1111111111
Db	318 GCCCTTGGCCCTCCAAAGTCAACGCCCGCATGGTGTACTTGCACTACCGAGGGTTCGA	259	1111111111	1111111111	1111111111	1111111111	1111111111
QY	233 ArgLeuArgGlnThrAlaLeuValAlaAlaIAlaIAlaIAlaIAlaIAlaIAlaIAlaI	252	1111111111	1111111111	1111111111	1111111111	1111111111
Db	258 CAGCTT-----GATGGTGGGTCCCGGAGCGCGGA-----GACGATCTG	220	1111111111	1111111111	1111111111	1111111111	1111111111
QY	253 SerIleIleArgAsnSerIleGlyLeuValAspValIleIle-----	266	1111111111	1111111111	1111111111	1111111111	1111111111
Db	219 TCGGTGTCAATACAGAAAGCGACGAGAGCGTCTCTTCTGTGGCCGACGAACGGCG	160	1111111111	1111111111	1111111111	1111111111	1111111111

TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99379253
PUBMED 10349636
REFERENCE

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Katsunari, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamakata, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stabili, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Balarelli, R., Barsh, G., Blake, J., Boffelli, D., Bult, C., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Boljuga, N., Fellench, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Welte, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawai, H., Kohsaki, S., and Hayashizaki, Y. Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 685-690 (2001)
MEDLINE 21085660
PUBMED 11217851
REFERENCE

5 (bases 1 to 3499)
Adachi, J., Aizawa, K., Akahira, S., Akiyama, T., Aono, H., Arai, A., Arahawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Horii, F., Hume, D., Inotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamuro, T., Yamashita, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and Hayashizaki, Y. Direct Submission
Submitted (18-ANG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Riken Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for

Further details:
cDNA library was prepared and sequenced in Mouse Genome Encyclopaedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTNN 3'], cDNA was prepared by using the reverse transcriptase-thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adaptor of sequence [5' GAGAGAGAGATCCAGAGTAAATTAATTAATTCCTCCCTCCCTCC 3']¹. cDNA was cleaved with BamHI and XhoI. cDNA of size comprised between 3 and 7 kb was selected before cloning. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI. Host: DH10B.
Location/Qualifiers

FEATURES

source

CDS

1..3499
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM:DB:4632435C11"
/db_xref="MGD:MGI:1912768"
/db_xref="taxon:10090"
/clone="4632435C11"
/tissue_type="skin"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
128..2422
/note="CARBOXYPEPTIDASE X2
data source: SPTT, source key: 054860, evidence: ISS
putative"
/protein_id="BAB31768.1"
/db_xref="GI:12859762"
/db_xref="MGD:MGI:1926006"
/gb_xref="MGD:MGI:1912768"
/note="MARITACCPALALALVALAVALAGYRAOAGAFEEPPDYISOELMR
RGRIYHPPPEPEDELEFSRMEHDLVEEOEOEPHOGHRIKPKAKRKKRKL
VETTPPGKNSNKKGRKSNLEKASDDHGGVPAHHDVSCPGLEFKLITPTDPLH
ASTSKRYGIGAHKRLNIOAGIENDFYOGAMCGNDLHOMLEVDAKRLITFTGYIT
OAGNSLIMLDWVTSYKVMVNSDSHTVTKNSGDMIFGNSKEKELPVNLELPPVPA
RYRINDQSMFNGSICRMEILGCPDPNNYHHRNEMTTDDLEPHNNKREKQ
LKKVNMCPNTRIVYNIKSHOGLKLYAVEISDHGEHEVGPFRHYTAGAGNGL
GRELLILHLHLOEYSAONARIVLVEERHILSLNPDGEKAYEGSGELGWSL
GRHTGDCIDNNPDLNLSLMEADDOONAPKVPVHYTAIPEMFSENNATVETRA
VIAMKRTIPVLGAGNCGEELVYAYPYDWRSLMKOEHPPDPDHPRLAVSYST
HRLMTDARRRVCTEDFOKEBGTVNGASHTVAGSLNDSYLTNCFELSIYVCGKY
PHSELPEEMENRESLIVFMEQVHKGIGIVADLOGKISNAVISEGVNNDIRAS
DGDYWRNLNGEYVTAKEGFTSTKNCWGYDMGATRCDFLTNTLNLRIMEET
FGKQPSLPSRLKLRGRKRGRC"
3479..3484
/note="putative"
3499
/note="putative"
polyA_site
polyA_signal
BASE COUNT 890 a 880 c 975 g 754 t
ORIGIN

Alignment Scores:

Pred. NO: 11.7 length: 3499
Score: 100.00 Matches: 148
Percent Similarity: 32.41% Conservative: 98
Best Local Similarity: 19.50% Mismatches: 279
Query Match: 2.12% Indels: 234
DB: 11 Gaps: 36

US-09-914-168-2 (1-919) x AK019509 (1-3499)

OY 31 GlnGlnAsnAnProAlaAsnIleIleAsnHsValProAlaHsAspTrhAlaIleAsn 50

DB 350 GAACGACGAGGACCGC-----CACGACGACGAGCGACACGAACT----- 385

OY 51 GlnAlaValAlaGlyAsnProValIleuLeuThrProGluGluGlnIleGlnAlaArgLeu 70

Dh 386 CCCAAGAGCCATCAGCCCAAGAGGCTCCCAAGAGGAGAGTTAGTTGCAGAGAC 445
Oy 71 AsnAlaIaIaGlyLeuAsnAlaLys-----ProGlnSerGlnAlaLeuAspVal 87
Dh 446 CCTCCACCCAGTAAATAATGACAACAGAAAGGAGAAAGCAAGAAATTGTGAGAAAGCT 505
Oy 88 AsnPheAspArgLysSer---ProIleSerArg-----IleGlyLysLysProPro 104
Dh 506 GCCAGTGAAGACCATGCTCCCTGCTGCTCATGAGAGATGTCAGAGAGATTGCCACCT 565
Oy 105 LeuGlyLeuAspMetSerValIleGluGluThrThrProLeuSerLeuGluIlePhe 124
Dh 566 CTTCGGCTCTG-----GAACATTAAATACACAGACTTCCACGCTGCAT 607
Oy 125 AlaGlnLysSer-----GlnLeuLeuGluGlnLeuThrValAsnMetGlyGlnAla 401
Dh 608 GGCTCCACATCGAAGCGTTATGGCCCTGGAGCCACCGGGGAGACTCAACATCCAGGCA 667
Oy 132 GlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyLysLysProAsnSerGluVal 151
Dh 668 GGCATTAAATGAATAATGACTTT-----TACGATGGGGCTTGGTGTCTGTAGGAAC 718
Oy 152 ValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyrAlaArg 171
Dh 719 GACTTGCAATCAGTGCATCGAAGGATGCGCGGCGCTGACCAAG-----TTCACAGGG 772
Oy 172 LeuPheAsnAspGlyValAsnLys-----ValProArgLeuLysAla 185
Dh 773 GTCATTATCCCAAGGAAGAACTCTCTCTGCTGAGTGCATGGGTGACATCTTAAGAATC 832
Oy 186 LysPheTyrGlnSerSerGlnSerGlnGlyLysThrSerAlaIleGlySer----- 201
Dh 833 ATGTGTAGCAATGACAGCCACACATGCTTACTGTGTAGAAATGAGATCTGGCGACATGANA 892
Oy 202 -----SerHisGlnLysThrGluProTyrAlaAsn-----IleLys 213
Dh 893 TTGTGAAGCAAAACAGTGCAGAGAGATTCCTGTGCTCAATAGAGCTGCCAGTCCCAATG 952
Oy 214 AlaAlaLeuGluAspIleThrGlnLysSerAlaMetAspLeuAsnGlySerIle----- 231
Dh 953 GCCCGCTACATTCGCATTAACCTCAGTCTGCTGTATGAT---AAGGGAGCATCTGCATG 1009
Oy 231 ----- 231
Dh 1010 AGGATGAGATCTTGGGCTGCCCACTGCCGATCCTAATATATTATCAGCGACCTAAT 1069
Oy 232 -----ProArgLeuArgGln 236
Dh 1070 GAGATGACCAACCAAGATACCTGGATTTTAAGCACCACTATTAAGAAATGGCCGAC 1129
Oy 237 ThrAlaLeuVal----- 240
Dh 1130 TTGATGAAGGTGTCAATGAATGTGCCCAATATTACAGAGATTTAACAACATTGGCANA 1189
Oy 241 AlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIleGly 260
Dh 1190 ACCCAACCAAGGCGCTGAATGTATGCGTGAAGATCTCT-----AACCATCTCTGG 1240
Oy 261 GluValAspValIleIleHisAspLeuGlyGluProValTyr-----IleAspTyr 277
Dh 1241 GAA-----CATGAAGTTGGTGAAGCCGAGATTCCACTACATCGAGGGGC 1285
Oy 278 ArgAlaValGluValArgGlyLys----- 285
Dh 1286 CACGGCAATGAGAGTCTGGAGAGAACTGCTGCTGCTGCATCTCTCTCCAG 1345
Oy 286 -----GlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuLeuIle 303
Dh 1346 GAATTAAGTGGGCGAGAACCCAGCATGTCCTGTGGTGGAGAGACTCGAATCCACAT 1405
Oy 304 GlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSer 323
Dh 1406 CTACCCCTCCCTCAATTCCTATGCTATGAG-----AAGGCTATGAGAGAGGTTCC 1456

Oy 324 AlaGlnHisGlyTyrPheAspGlyArgTyrProLeuAspArgSerValAspValIleLeuPro 343
Dh 1457 GAGTTGGAGGCGGTGGTCCCTGGGAGCTGTGACCCATGATGGCATGCATATC-----AAC 1510
Oy 344 AspAsnThrAlaAspVal-----SerLeuIleTyrAspThrGlyThrLysIleTyrArgPheAsp 362
Dh 1511 AACCACTTCGGAATTTAAATCTCGCTGTGGAGGACAGAGACAG----- 1558
Oy 363 GluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProAspLys 382
Dh 1559 -----CAGATGCCCAACGAG 1576
Oy 383 LeuProValLysArg---GluLeuLeuGlnGlnLeuThrValAsnMetGlyGlnAla 401
Dh 1577 GTCCCAACACACTACATTCGCATCCCTGAGTGATTTCTGTCTGAATGCAACAGTGGCC 1636
Oy 402 TyrAsnLeuGlnAlaValArgAla----- 409
Dh 1637 ACAGAGACAGAGCCCTCATCGCTGATGAGAGATCCCGTTTGTCTGGAGGCAAC 1696
Oy 410 ---LeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleVal 428
Dh 1697 CTACAGGGGGGTGAGCTGTGCTGCGCATACCCCTATGACATGCTGCGGTCCCTGTGAAG 1756
Oy 429 PheProGluArgGluGlnIleGlnAsnAspGlnVal-----SerPheGlnGln 444
Dh 1757 ACCCAGAGACACACCCCAACACTGATGATCATGTCTCCGCTGGCTGGGTATTCCTAC 1816
Oy 445 SerSerSerSerArg-----ThrGluProAla 453
Dh 1817 GCCTCCACTCACCGCCTCATGACAGATGCCAGAGGCGAGTGTGCCACAGGAATTTT 1876
Oy 454 GlnValAspGlySerThrLeuGluProValIleGluThrValGluLeuThrArgGlyIle 473
Dh 1877 CAGAGAGAGAGGCGCCCTCAATGGGCT-----TCTGGAGCACAGTGGTGAAGT 1930
Oy 474 LeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsn 493
Dh 1931 CTAAAGCATTTTCAGC-----TACCTGCATACAACTGCTTTGAGCTGTGCATCTAC 1981
Oy 494 LeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAsp-----Arg 509
Dh 1982 GTGGGCTGTGATTAATACCCACGAGACGACCTCCGAGGAGGAGGAATAACCCG 2041
Oy 510 ValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAsp 529
Dh 2042 GAGTCTCTGATTTGTTCATGAGACAGGTTTCATCGAGCATCAAGGCAATAGTGAAGAT 2101
Oy 530 AlaVal---SerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIle 548
Dh 2102 TTACAGGGAAGAGATTTCAATGCTGCATCTGTGGAAGGTGTTAACCATGACATC 2161
Oy 549 AspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSer 568
Dh 2162 -----CGGACAGCC-----AGCGATGGGATTAATCTGCCTCA 2194
Oy 569 LysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyIleIleGly 588
Dh 2195 CTGAACCCGTGGCAATATGTGTGCACACGCCAAGCGCGAA----- 2233
Oy 589 LeuGlyTyrPheLysSerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIle 608
Dh 2234 ---GGCTTATACCTTCACCAAGAACTGATGCTT-----GluLeuArg 620
Oy 609 AsnArgAspGlyTyrGlnIleGlyAla-----GluLeuArg 620
Dh 2267 -----GGCTATGATTTGGAGCTACTCGGTGTGACTTACCCCTCAACAAAGCAAC 2317
Oy 621 LeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisPro 639
Dh 2318 CTGGCTAGGATTAAGAAATTTATGAGACATTTTGGAGACAGCAGCTCTCAGCGCTACCC 2374

[illegible]

Qy	136	AsnaasyrTlRleprogluTyGlnclYgluGlnInProAansSergluVal-----Val	152
Db	440	TCA-----CCAGTTTACGAGCGGGAACCTCAAGTTAGTTGAGAGCATT	490
Qy	153	ValPro 154	
Db	491	GTGCCA 496	
RESULT 38			
LOCUS	AY108441	1514 bp	linear
DEFINITION	Zea mays	PC0083558	mRNA sequence.
ACCESSION	AY108441		
VERSION	AY108441.1	GI:21211519	
KEYWORDS	HTC.		
SOURCE	Zea mays.		
ORGANISM	Zea mays.		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	1 (bases 1 to 1514)		
TITLE	Hainey, C.F., Dolan, M., Miao, G. H., Vogel, J.M., Whitesitt, M.S., Arthur, L.W., Hanfey, M., Morgante, M. and Tingey, S.V.		
JOURNAL	Maize Mapping Project/Dupont Consensus Sequences for Design of Overgo Probes		
REFERENCE	Unpublished (2002)		
AUTHORS	2 (bases 1 to 1514)		
TITLE	CoE, E.C.		
JOURNAL	Direct Submission		
REFERENCE	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA		
FEATURES	Location/Qualifiers		
source	1..1514		
	/organism="Zea mays"		
	/db_xref="maizedb:634674"		
	/db_xref="taxon:4577"		
	/clone="PC0083558"		
	/clone_lib="Maize Mapping Project/Dupont Consensus Library"		
	/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"		
BASE COUNT	452 a	323 c	371 g
ORIGIN	368 t		
Alignment Scores:			
Pred. No.:	3..05	Length:	1514
Score:	99.50	Matches:	74
Percent Similarity:	35.48%	Conservative:	64
Best Local Similarity:	19.02%	Mismatches:	146
Query Match:	2..10%	Indels:	105
DB:	11	Gaps:	15
US-09-914-168-2 (1-919) x	AY108441 (1-1514)		
Qy	193	SergluThrsalalaleglsrSerHislnls-----	205
Db	65	TCAAGCGACGCGATGCGGATGCGCGCGCTCGACGCGAGGAGCGCCCTCATCTCTCC	124
Qy	206	---ThrgluProTyraAsnlelYalAlaAlaLengluAspIleThrglnlSera	244
Db	125	TTCACCGCTCCCAACCGCTCTGTGGCGCGCTTCCCTCTGCGTGAACGATCTCTCGCC	184
Qy	225	MetaspleuansnglSerleProArgLeuArglnlThralaleuValAlaAlaArgAla	244
Db	185	GTCTCGACACGAGATGCGTCCA-----	208
Qy	245	ValglYtrYrAspIleAspIleSerllelelAargAsn-----SerlelgYglu	261
Db	209	CTGTGTCGACGCTCTATTCTCACTCAATCAATCTGTGAGAACCGGATGACAGTGTACGAGAT	268

TITLE
JOURNAL
Direct Submission
Submitted (10-JUL-2000) Yoshihide Hayashizaki. The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
Please visit our web site (http://genome.gsc.riken.go.jp/) for
further details.
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in RIKEN.
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
15' GAGAGAGAGAGCGCCGCACTGACTTTTCTTTTCTTNN 3'. cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through one round of normalization to Rot = 5.0. Second strand cDNA
was prepared with the primer adapter of sequence 5'
GAGAGAGAGAGAGATCCAGAGCTCAATTATTTAATTAACCCGCCCC 3'. cDNA was
cleaved with XhoI and SclI. Cloning sites, 5' end: SclI; 3' end:
XhoI. Host: SOLR.
Location/Qualifiers
1. 1562
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:0610009009"
/db_xref="MGI:1909738"
/db_xref="taxon:10090"
/clone="0610009009"
/sex="male"
/tissue="kidney"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1. 1562
/gene="D15wsu77e"
50. 1423
/gene="D15wsu77e"
/note="DNA segment, Chr 15, Wayne State University 77,
expressed
data source: MCD, source key: MGI:106312, evidence: ISS
putative"
/codon_start="1"
/protein_id="BAB2209.1"
/db_xref="GI:12832417"
/db_xref="MGI:106312"
/translation="MSIHFSRSSTAYPGKAGVRLSSGRASGRSLYGLSSRPVA
VRSAYGPGVAGCIREITINOSLLAPLSDVIDPTTOVROEEREQIKTLNNKRSLEK
VRFLOONKMLETKMALLOEOKSAKSSQDPRLEAOIAGLRQOLETLQDGRSEVEL
RMOVDVEDFKNKYEEINRRTAENEVELLRKDYDAATNVELEAKDSQDEIN
LKTTHETELAESOIPTSVYLSMDSNLSLDLDIADVKQYEBMAHDSRAEAM
VOTKRETLAOGAKHGDDLRNREIAENRSIORLOAEITLKNORAKLESIAE
EDGELAIIDAHAKOGELEALOKAKODVROLREYOLELTLAIDIEATRYKLEEG
EESRLSGDGMFVNISVNSTGNGSKLIFGGTMSNLSFSGGALRAYSIKTTST
TRRGTHN"
BASE COUNT 367 a 427 c 482 g 286 t
ORIGIN
Alignment Scores:
Pred. No.: 3 23 Length: 1562
Score: 99.50 Matches: 100
Percent Similarity: 33.33% Conservative: 63
Best Local Similarity: 20.45% Mismatches: 163
Query Match: 2.10% Indels: 163
DB: 11 Gaps: 21
US-09-914-168-2 (1-919) x AK002434 (1-1562)
QY 379 AspProAspIysLeuProValIyArGluLeuLeuGluInLeuLeuThrValAsnMet 398
|||||
Db 272 GATCCACCATTCACGACGCGCCAGAGAGAGAGGAGATCAAGACCCCTCAACAC 331
|||||

QY 399 GlyIuAlaIyArAsnLeuGluInAlaValArgAlaLeu-----SerAsnAspLeuIleAla 416
|||
Db 332 AAATTCGCGCTCTTCATCATGACAGAGTACGCTTCCTGGAAACGACACACAGATGCTGGAG 391
QY 417 ThrArgIyPheAsnMetValAsnThrGluIleValPheProGluArgGluInIleGln 436
|||||
Db 392 ACCAAGTGG-----GCGCTGCTG 409
QY 437 AsnAspGlnValSerPheGluInSerSerSerArgThrGluProAlaGlnValAsp 456
|||||
Db 410 CAGGAACAGAACTCGCCACAGACGACCCAGCTCCCGCAATCTTGAGGCTCAGATT--- 466
QY 457 GluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetSrp 476
|||
Db 467 ---GCTGGGCTGAGGACAGAGCTGAGACACTGAGCTGGATGGGCGCGCTG----- 517
QY 477 IleSerProIleGluPheSerIleAsnLeuIleGlnAspIysLeuAsnLeuValAla 496
518 -----GAG 520
QY 497 AlaIyAlaArgHisLeuIyAspMetProAspArgValIleAlaIleAsnHisAsp 516
|||
Db 521 GTGGAACCTCGGACATCGAGATGCTGGAAGAT-----TTCAAGAATGAATGAA 574
QY 517 AspGlyValAsnArgSerIleLeuGlyIyArgIleSerAspAlaValSerAlaValArg 536
|||
Db 575 GAGGAGATCAACCGACGC----- 592
QY 537 AlaIleLeuProAspIySerGluAsnGluValIleAspLeuProGluArgThrAlaLeu 556
|||
Db 593 -----ACGCGCTCGAGATGAGTTGTGTTCTG----- 622
QY 557 AlaAsnArgIySThrProAlaAspValIyGlnSerIySLeuValProLeuIyValPhe 576
|||
Db 623 -----AAGAGATGTGATCGCGCTACACACAGAGCTGAGCTGGAG----- 667
QY 577 ValAlaSerAspIySProArgAspGlyGlnIleGlyLeuGlyIyPheIySerAspThrGly 596
|||
Db 668 GCCAAGCGGACAGCTGACGAGT---GAGATCACTC----- 703
QY 597 ThrArgLeuValThrIySPhelGluHisAsnLeuIleAsnArgAspIyIyGlnAlaGly 616
|||
Db 704 -----CTCAGACCTTCACGACGACAGATVA----- 730
QY 617 AlaGluLeuArgLeuSerGluAspIySLeuGlyValIyLeuIyValThrIySProLeu 636
|||||
Db 731 GCAGAGCTGACAGTCGACATCTCAGACATCTGTGTGCTGTCATGCAACACAGCCGC 790
QY 637 SerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyIyGlnGlnGluValPheGly 656
|||
Db 791 TCCCTGACTTGGATGTATCATTTGCTGACGCTCAAGCCACGATGAGGATGGCCAC 850
QY 657 HisSer----- 658
|||
Db 851 CACAGCCGCGACAGGCTCAAGCTGTGTACACAGCAAGTTTGAACACTCCAGCCCCAG 910
QY 659 -----ThrAsnGlyPheAspLeu---SerThrArgThrLeuGlnHisGlnIleSerArg 675
|||
Db 911 GCTGGAGACGAGGGGATGACCTCCGACACACCGGAAATGAGATTGGGAGATGAAACCCG 970
QY 676 SerIleIleGlnAsnGlyIyTrpAsnArgThrIySerLeuArgIyArgLeuAspIyS 695
|||||
Db 971 TCTATC-----CAGAGGCTGACGAGGATTTGACACC 1003
QY 696 LeuIySThrGln-----AlaProProGluThrIyPpGlnAsp 707
|||||
Db 1004 TTGAAGAACCCAGCGCTCCCAAGTTAGACTCCAGATCCAGATCCAGAGGCTGAACAAGGGGAG 1063
QY 708 LeuProValAspPheValAsnGlyIySProSerGln---GluAlaLeuLeuAlaGlyVal 726
|||
Db 1064 CTGGCAATCAAGATGCTCATGCTAACACAGGGAGAGCTGGAGCTGGCCCTG----- 1114
QY 727 AlaValHisIySThrValAlaAspAsnLeuValAsnProMetArgGlyIyArgGlnArg 746

Db	1115	-----CAGAAAGGCCAACGACGATGTGGCCCGCACCACTTCGAGACTACAG-----	1159
Oy	747	TyrSerLeuGIValIGlySerSercIleuValISerAsnAlaIsnMetAlaISlleAlaArg	766
Db	1160	-----GAGCTCCTGAACACCAGGCTGCAGCGCTGACCATTTGAGATTCGCCACCTTACGC	1210
Oy	767	AlaGlyIleSerGlyValTyrrSer-----PheGlyAsp-----	777
Db	1211	AAGCTCGTGAGCGGCGAGACAGCGCGTTGTCITGTGAGACGGAAGAATGACACTGTCAATATC	1270
Oy	778	-----AsnAlaTyrrGlySerAsnAlaAlahisGlnMetThrGlyGlyTlEglnAla	794
Db	1271	TCTGTGTGAATTCCACATCGGGGGGCAATGTGGCAAGCTTACTTCCTGGGAGAACCATGGGA	1330
Oy	795	GlyTyrrIleTrpSerAspAsnPhesnhIstValProTyrrATgLeuArgPhePheAlaGly	814
Db	1331	AGCAATGCTCTGAGC-----TTCACTGTGGG	1354
Oy	815	GlyAspGlnSerIleArgGlyTrrAla	823
RESULT 40		: : :	
LOCUS	W29713/c	W29713	1022 bp mRNA linear EST 11-SEP-1996
DEFINITION		mco5a09.f1 Soares mouse p3NMF19.5 Mus musculus cDNA clone IMAGE:338008 5' similar to gb:V00714 Mouse gene for alpha-globin (MOUSE);, mRNA sequence.	
ACCESSION		W29713	
VERSION		W29713.1	GI:1309862
KEYWORDS		EST.	
SOURCE		house mouse.	
ORGANISM		Mus musculus	
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS		1 (bases 1 to 1022)	
		Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Gelsel,S., Kucada,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.	
TITLE		The Mashu-HMI Mouse EST Project	
JOURNAL		Unpublished (1996)	
COMMENT		Contact: Marra M/Mouse EST Project WashU-HMI Mouse EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel.: 314 286 1800 Fax: 314 286 1810 Email: mouseest@watson.wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MG1:219408	
FEATURES		Seq primer: ETPrimer	
SOURCE		High quality sequence stop: 350.	
		Location/Qualifiers	
		1..1022	
		/organism="Mus musculus"	
		/db_xref="taxon:10090"	
		/clone="IMAGE:338008"	
		/cdate="11-Sep-1996"	
		/dev_stage="19.5 dpc total fetus"	
		/lab_host="DH10B (ampicillin resistant)"	
		/note="Vector: pT7TS3 (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; Ist strand cDNA was primed with a Not I - oligo(dT) primer [5', TGTTACCAATCTGCAAGTGGGAGCGCGCATTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7TS3 vector (Pharmacia). Library was through one round of normalization to a Cot = 5. Library constructed by Bento	

Soares and M. Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."

Alignment Scores:	
Pred. No.:	1.75
Score:	99.00
Percent Similarity:	38.74%
Best Local Similarity:	24.02%
Query Match:	2.09%
DB:	14
	Gaps: 19

```

OY      30  AAlngInglAsnAspProhlaSnIleIleAsnHis-ValProAla-----Hisas  46
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      923  AGCAAGAGACACACCACATCCATACCCGCGATACACTCTCCCGCCAGCATATCA  864
OY      46  pThAlaIleasngInAlaLyAlaGlyAsnProProValIleuLeuThrProGluGlnI  66
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      863  CCGTTTATCTCGATTAACACAAA-----AACCCCGCTTTTAACCCGCAAAAAGGCGCG  810
OY      66  eGlnAlaIleArgLeuAsnAlaIaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspVa  86
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      809  ACAATTCGCCCATTT-----CCAGCGCAAAATTTCTTG-----  779
OY      86  lValAsnPhaeAspGlnSerProIleSerArgIleGlyGluGlnSerProProLeuGlnI  106
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      778  -----GAAAGGCCAATC--CGTGCGG-----  758
OY      106  yLeuAspMetSerValIleGluGlnThrThrProLeuSerLeuGluGlnLeuPheAlaG  126
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      757  -----CTTTTCCGATATCTCCCGCACCTCTCCAAAAGGGAGATCTGCTCA  711
OY      126  nGluSerThrGluMetGlyIleAsnProAsnAspTyr-----IleProGluThrGlnG  144
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      710  AGCGGAGTTAAATTCGGG--AACCCACAGAGTTTTCACAGTACCGCAGTTTGAAA  654
OY      144  yGluGlnProAsnSerGlnValValValProProThrLeuGluProGluLys-----  161
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      653  ACACACCCCAATCCCAACTAAATTAATTAACCCCTACACTAAAGAGATMAAGCTTTCCG  594
OY      162  ProGlyLeuIleLysArgLeuTyr-----  169
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      593  CCCATGTTATTTTTCATCCCTCTCCGCCACTTCCTCTCATCTTTTNTCTCTAC  534
OY      170  -AlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheArg  189
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      533  TCAGGCTTATTTCAAGACATGATGATACAGTTGCAAGCAGAGAAAGAACGGCATCCCA  474
OY      189  nSerSerGlnSerGlyGluThrSer-----AlaIleGlySerSerHisGln-----  204
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      473  GAATGCACCCCGCAGAACGCGATTACGGTACTTGGAGTCCACACGCGTCTCACAG  414
OY      205  -----LysThrGluProGlyrAlaAsnIleLysAlaAla-----Le  216
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      413  AGGCAAGAAATTTATCCAGAAGAGCGCATGCAACCGGGGGTGAATTCGGCAGGGTGCGCT  354
OY      216  uGluAspIleThrGlnGlnUser-----AlaMetAspLeuAsnGlySerIleProAr  233
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      353  AGCCAAAGTCCACAGCAGGCGAGTGGCTCAGAGAGCTTGAAGTTACGGGATCCACAGCAG  294
OY      233  gLeu---ArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLe  252
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      293  CTTGTGGCGATCGAGTCTCTTCAGACACACAGGGCGACCGGAGGCTCATCGAG-GTGGC  235
OY      252  uSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGluPro  272
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      234  CTGCAGCATTTGGCCACAGCATCGGAGCACTTCTTGGCGTGAACCCCTTGACCTGGGACAGC  175
OY      272  oVal---TyrIleAspTyrArgAlaValGlnValArgGlyGluGlyAlaAsp-----  288
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Db 174 CGTGGCTTACATCAAGTAGGAGAGTAGTCTTGTGGTGGGAGCTAGCAACATCC 115
QY 289 -----AspLysAlaPheThrValAlaAspGluValPro-----Le 301
Db 114 TTTCAGGGCTTCAGCTCCATATTCAGCACCATGGCCACCAATCTTCCCCAGGAGCCT 55
QY 301 uLeuIleGlyAspValPhe-----HisHisGly 310
| : : : | | | | |
54 TGATGTGCTTTTGTCTCTCCCGAGAGACCATGCT 18

Search completed: May 6, 2003, 01:45:34
Job time : 2095 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 6, 2003, 01:32:51 ; Search time 669 Seconds
(without alignments)
6472.581 Million cell updates/sec

Title: US-09-914-168-2
Perfect score: 4727
Sequence: 1 MSKPVLFANRSEMPVALAAY.....TGCKEENGRKILHFIFGPF 919

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 6113114 seqs, 2355906770 residues
Total number of hits satisfying chosen parameters: 12226228

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_model -DEV=xlh
-O/cgcn2.1/USPTO.spool/US09914168.rnpuat_28042003_151617_6869/app_query.fasta.1.1095
-DB=Pending_Patents_NA_New -QPMF=fastap -SUFFIX=rpnp -MIMATCH=0.1 -DOOPCT=0
-DOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=40
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09914168.@CGN1.1.465.@runat_28042003_151617_6869 -NCPU=6 -ICPU=3
-NO_XMPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Pending_Patents_NA_New:

1: /cgcn2_6/ptodata/2/pna/PCT_NEW_COMB.seq:*
2: /cgcn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgcn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgcn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgcn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgcn2_6/ptodata/2/pna/US09_NEW_COMB.seq2:*
7: /cgcn2_6/ptodata/2/pna/US09_NEW_COMB.seq3:*
8: /cgcn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
9: /cgcn2_6/ptodata/2/pna/US10_NEW_COMB.seq2:*
10: /cgcn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*
11: /cgcn2_6/ptodata/2/pna/US60_NEW_COMB.seq2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	620.5	13.1	1818	8	US-10-419-128-14784
2	620.5	13.1	1818	9	US-10-366-683-14784
3	500.5	10.6	1860	8	US-10-417-886-3279
4	436.5	9.2	1737	6	US-09-913-101-1
5	427.5	9.0	1731	6	US-09-913-101-3
6	405	8.6	4497	8	US-10-419-128-14663
7	405	8.6	4497	9	US-10-366-683-14663

C	8	362	7.7	726	8	US-10-419-128-14966	Sequence 14966, A
C	9	362	7.7	726	9	US-10-366-683-14966	Sequence 14966, A
	10	353	7.5	1416	9	US-10-144-771-21979	Sequence 21979, A
	11	351	7.4	1365	10	US-60-453-134-81	Sequence 81, App1
	12	319	6.7	944	8	US-10-399-416-9	Sequence 9, App1
	13	265.5	5.6	2379	9	US-10-148-534-6	Sequence 6, App1
	14	259	5.5	2391	9	US-10-320-800-5	Sequence 5, App1
	15	259	5.5	2394	9	US-10-148-534-2	Sequence 2, App1
	16	259	5.5	2394	9	US-10-148-534-10	Sequence 10, App1
	17	238.5	5.0	6617	7	US-09-689-065A-1	Sequence 1, App1
	18	238.5	5.0	6617	7	US-09-689-065B-1	Sequence 1, App1
	19	206.5	4.4	3012	6	US-09-988-067B-7	Sequence 7, App1
	20	200	4.2	2778	8	US-10-335-977-157	Sequence 157, App
	21	186	3.9	2466	8	US-10-419-128-4442	Sequence 4442, Ap
	22	186	3.9	2466	8	US-10-366-683-4442	Sequence 4442, Ap
	23	174.5	3.7	516	8	US-10-419-128-14662	Sequence 14662, A
	24	174.5	3.7	516	8	US-10-366-683-14662	Sequence 14662, A
	25	162.5	3.4	1530	8	US-10-419-128-4165	Sequence 4165, Ap
	26	162.5	3.4	1530	9	US-10-366-683-4165	Sequence 4165, Ap
	27	160.5	3.4	2163	9	US-10-282-122A-33708	Sequence 33708, A
	28	158	3.3	1812	9	US-10-010-160-15	Sequence 15, App1
	29	157	3.3	1890	8	US-10-369-493-42248	Sequence 42248, A
	30	154	3.3	2349	10	US-60-453-134-135	Sequence 135, App
	31	154	3.3	3111	9	US-10-282-122A-27975	Sequence 27975, A
	32	152	3.2	2183	8	US-10-369-493-44808	Sequence 44808, A
	33	150	3.2	2763	8	US-10-282-122A-33378	Sequence 33378, A
	34	148.5	3.1	4802	9	US-10-263-929-30	Sequence 30, App1
	35	148.5	3.1	4802	9	US-10-267-502-24	Sequence 24, App1
	36	146	3.1	6609	9	US-10-172-502-1	Sequence 1, App1
	37	145.5	3.1	7542	9	US-10-320-800-39	Sequence 39, App1
	38	145.5	3.1	7545	9	US-10-282-122A-29937	Sequence 29937, A
	39	145.5	3.1	2272325	9	US-10-018-470A-1	Sequence 1, App1
	40	145.5	3.0	4473	9	US-10-282-122A-29386	Sequence 29386, A
	41	143.5	3.0	3920	9	US-10-282-122A-14980	Sequence 14980, A
	42	142.5	3.0	4293	9	US-10-282-122A-42267	Sequence 42267, A
	43	142.5	3.0	1752	8	US-10-369-493-23985	Sequence 23985, A
	44	142	3.0	1752	8	US-10-282-122A-37677	Sequence 37677, A
	45	141.5	3.0	5631	9		

ALIGNMENTS

RESULT 1
US-10-419-128-14784
; Sequence 14784, Application US/10419128
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/10/419,128
; CURRENT FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US/09/252,991
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ. ID NOS: 33142
; SEQ ID NO 14784
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-419-128-14784

Alignment Scores:

Pred. No.: 8.39e-41
Score: 620.50
Percent Similarity: 40.61%
Best Local Similarity: 24.45%
Query Match: 13.13%
DB: 8
Length: 1818
Matches: 177
Conservative: 117
Mismatch: 225
Indels: 205
Gaps: 15

```

US-09-914-168-2 (1-919) x US-10-419-128-14784 (1-1818)
OY 210 Alaasn1lelyala1aleuengluaSPIlethnglIngluSerAlaMetaspLeuansngly 229
    |||||
Db 199 GCCAATATTCGAGCGCCCTAGCTGGTACCGCGGAC-----GAGGCG 246
OY 230 Ser1leProArLeuArGlnThrAlaLeuVal-----AlaAlaArGAlaVal 245
    |||||
Db 247 GCGTTGCCAAGCTTTCGTCGCAATGCGGAGCGCGGAGGAAAGCGCGCCCGCGCGCTC 306
OY 246 GlyTyrTyrAsp1leAspLeuSer1leIeArGAsnSer1leGlyGluValAspVal1le 265
    |||||
Db 307 GCGTACTTCCAG-----GCCCAATGCACAGCGAG 336
OY 266 1leHisAsp-----LeuGlyGluProValTyr 274
    |||||
Db 337 GTGAAGACGCGGACGCGCGGAGCTCAAGCTCAAGTAGTCCCGGAGCGCTGGCGC 396
OY 275 1leAspTyrArGAlaValAlaGluValArGlyGluGlyAlaAspAspLysAlaPheThrThr 294
    |||||
Db 397 CTGCGCCAGGTCAACATCCAGGTGCTCGGAGCGCGGAGCTGGAGAGCTTCGCGCTTG 456
OY 295 ValAlaAspGluValArProLeuLeu1leGlyAspVal1PheHisGlyLysTyrGluThr 314
    |||||
Db 457 CCGTCGCGGACAG---CAACTGAAGCGGCGCGGAGCTGAACAGGAGGTCTACGAGGAC 513
OY 315 LysLysAsnLeu1leGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArGTrPLeu 334
    |||||
Db 514 GCCAAGCGCGGTATCCACACAGCGCTTCGCGGAGCTTCTTCAGCGCGCTTCGAGC 573
OY 335 AspArgSerValAspVal1leLeuProAspAsnThrAlaAspValSer1leIeTyrAsp 354
    |||||
Db 574 ACCGAGCGCCCTGACATCGATCCCGCGCGCGCATCGCGCATCGATCGATCGATCGAC 633
OY 355 ThrGlyThrGlnTyrArGpPheAspGluValArPhePheThr1leAspProLysThrAsn 374
    |||||
Db 634 AGCGCGCAGCGCTACCTCGCTCGGACAGTGCAGTTC----- 669
OY 375 GlnLeuThrThrAspProAspLysLeuProValLysArGlnLeuLeuGluGlnLeuLeu 394
    |||||
Db 670 -----GACGCGGACTCGATC---ATCGAGAGAGAACTGTCGCGCGCTGCTG 714
OY 395 ThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArGAlaLeuSerAspLeu 414
    |||||
Db 715 CCGTTCAAGCGCGGCAACCTTAGTACTCGAAGTATCGCGGAGCTCAACACAGACTG 774
OY 415 1leAlaThrArGTrPheAsnMetValAsnThrGlu1leValPheProGluArGluGln 434
    |||||
Db 775 CAATCCAGCGGCTACTTC----- 792
OY 435 1leGlnAsnAspGlnValSerPheGluGlnSerSerSerArGTrhGluProAlaGln 454
    |||||
Db 792 ----- 792
OY 455 ValAspGluSerThrLeuGluProVal1leGluThrValGluLeuThrAspGly1leLeu 474
    |||||
Db 793 -----GAGGAGAGTGGCG 804
OY 475 MetAsp1leSerPro1leGluPheSerAlaSerAsnLeu1leGlnAspLysLeuAsnLeu 494
    |||||
Db 805 GTGAGCGCGCGCGGAGCCAGGCTCAGGCC----- 834
OY 495 ValAlaAlaLysAlaArGlnHisLeuTyrAspMetProAspAspArGValLeuAla1leAsn 514
    |||||
Db 834 ----- 834
OY 515 HisAspAspGlyValAsnArGSer1leLeuGlyArG1leSerAspAlaValArAlaVal 534
    |||||
Db 835 -----GACGCGCGCGCGCGAGCGGATC----- 855
OY 535 AlaArGAla1leLeuProAspGluSerGluAsnGluVal1leAspLeuProGluArGTrh 554
    |||||
Db 855 ----- 855

```

```

OY 555 AlaLeuAlaAsnArGlySerThrProAlaAspValTyrGlnSerLysValArProLeuTyr 574
    |||||
Db 856 -----CCGGTACCG 864
OY 575 ValPheValAlaSerAspLysProArGAspGlyGln1leGlyLeuGlyTyrGlySerAsp 594
    |||||
Db 865 CTCGCGCTGGAGAGCGCGGAGCGGAGCATGGCGGCTCGGCGCTCGGCTTCGACAGCAG 924
OY 595 ThrGlyThrArGLeuValThrLysPheGluHisAsnLeu1leAsnArGAspGlyTyrGln 614
    |||||
Db 925 CTCGCGCGCGCGCGGAGGTCACTGAGACCGCGCATGGGTAATGCCAGAGACATAGC 984
OY 615 AlaGlyAlaGluLeuArGLeuSerGluAspLysGlyVal1LysLeuTyrAlaThrLys 634
    |||||
Db 985 CTGGATTCCTCGGAATCTCCGACACAGGAGGAGAACTCGCGGCTGTACAGATT 1044
OY 635 ProLeuSerHisProLeuAsnAspGlnLeuArGAlaThrLeuGlyTyrGlnGlnGluVal 654
    |||||
Db 1045 CCCCTCGATCCGCGGCTCACGAGACAGCTGCGCTTCACAGCGGCTACAGCTTCGAA--- 1101
OY 655 PheGlyHisSerThrAsnGlyPheAspLeuSerThrArGThrLeuGluHisGlyLysSer 674
    |||||
Db 1102 -----GACTAGTCGACACCGAGAGCAAGCTGACGCTTGGCGCGGAA 1146
OY 675 ArgSer1leGlnAsnGlyTyrPAsnArGTrhTyrSerLeuArGTrArGLeuAsp 694
    |||||
Db 1147 TGGCAGACAGAGCGCGCGATGCTGGACAGCGGTGCTTCGGAAGCTGATCGCGGAG 1206
OY 695 LysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspThreValAsn 714
    |||||
Db 1207 GAGTACAAAGCTG----- 1218
OY 715 GlyLysProSer-----GlnGluAlaLeuLeuAlaGlyAlaValAlaHisLysThrVal 732
    |||||
Db 1219 GGGCAGCACTCCGCGGCTGAGCACTTCCTGATGCGGGATTCGCTACTGCTGAGAG 1278
OY 733 AlaAspAsnLeuValAsnProMetArGGLTyrArGlnGlnArGTrSerLeuGluValAlaGly 752
    |||||
Db 1279 ACCGACAAAGAGTCAGACCCGACCATGGTACCGCTGAGATTCAACGCAAGGGGCGG 1338
OY 753 SerSerGlyLeuValSerAspAlaAsnMetAla1leAlaArGAlaGlyLysSerGlyVal 772
    |||||
Db 1339 AAGAGAGCGCTGTGGCGCGAGCTGCTTCATATGTCAGCTAGGCAAGGGCGCTG 1398
OY 773 TyrSerPheGlyAspAsnAlaTyrGlySerAsnArGAlaHisGlnMetThrGlyLys 792
    |||||
Db 1399 ACCAGCTTCGCGCGC-----GCCATCGCTGCTGCGCGCTG 1437
OY 793 GlnAlaGlyTyr1leTrpSerAspAsnPheAsnHisValArProTyrArGLeuArGpPhe 812
    |||||
Db 1438 CAGGTAGCGGGAATCCGCCAACAGCTACAGTCAATCCCGCGCTGCGCTTC 1497
OY 813 AlaGlyLysAspGlnSer1leArGlyTyrAlaHisAspSerLeuSerPro1leSerAsp 832
    |||||
Db 1498 GCTGGCGGAGACAGAGCTGCGTGCATCGACTCGGAGCTGCGCGCTGCGGAGAAATCC 1557
OY 833 LysGlyTyrLeuThrGlyGlnVal1leAlaValAlaGlyThrAlaGluThrAsnTyrGlu 852
    |||||
Db 1558 GATGGGAGAAAGATCGCGCGCGCTACATGATCCCGGAGCGTGAAGATCAATATCCG 1617
OY 853 PheMetLysAspLeuArGLeuAlaValPheGlyAsp1leGlyAsnAlaTyrAspLysGly 872
    |||||
Db 1618 CTGGCGGAGCGCTGCGCGCTGCGACTTCGTCGACAGGCGGACGCTTCACACTCGCTG 1677
OY 873 PheThrAsnAspThrLys1leGlyAlaGlyValAlaGlyTyrPalaSerProValAla 892
    |||||
Db 1678 GACTTCCCTCGATCAAGACCGGCTGCGCTTCGCGCGCTGCGCTGCGCTGCGG 1737
OY 893 GlnValArGValAspValAlaThrGlyValLysGluGluGlyAsnPro1leLysLeuHis 912
    |||||
Db 1738 CCGTTGGCGCTGACCTGGCCCATGCGCTCGACAGACAGCGGCGT---TTCGCGCTGAC 1794

```

Oy	210	AlaasnlllelysaAlaAlaLeuclnAspIleThrgIngluSerAlaMetaspleAasnGly	229
Db	199	GCCATTATCGAGGGCCCTACGTAGCGTAGCCTCGGCAAGCGGAC-----GAGCGC	246
Oy	230	SerIleProAArgLeuArgGlnThrAlaLeuVal-----AlaAlaArgAlaVal	245
Db	247	GGCTTGGCAACGCTTTCGGTCCGAAITGCCAGAGCGCAGCGGAAAGAGCGGCGCCAGGCCCTC	306
Oy	246	GlyTyrTyrAspIleAspLeuSerIleLeaArgasnSerIleGlyGluValAspValIle	265
Db	307	GGCTACTCTCCAG-----GCCAGATGACAGCAGCGAG	336
Oy	266	IleHisAsp-----LeuGlyGluProValTyr	274
Db	337	GTGAGAGCGGCGCAAGCCGCCAGACCTCACGCTCAAGTAGTGGCCGGCGAGCCCGTCGCGC	396
Oy	275	IleAspTyrArgAlaValAlaGluValArgGlyGluGlyAlaAspAspIlyAlaPheThrThr	294
Db	337	CTGGCCCGAGGTGAACATCCAGGTGCTGGCGCAGGCGCGCAGCCTGGAGAGCTTCCGCTTG	456
Oy	295	ValAlaAspGluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThr	314
Db	457	CCGTGGGGCGAG---CAACTGAAGGGGGCGCCGGAAGCTGAACCAAGGGGTGTCTACAGAGGAC	513
Oy	315	LysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrLeu	334
Db	514	GCCAAAGCGCGGTATCCAGAACCAAGGCTCGCGGTACGGCTTCTCCAGGCGCGCTTACGC	573
Oy	335	AspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAsp	354
Db	574	ACCACGCGCCGTAGCATGATGATCCGCGCGCGGATCGCGCGCATCGACTGGTCTACGAC	633
Oy	355	ThrGlyThrGlnTyrArgPheAspGluValAlaPhePheThrIleLeaProLysThrAsn	374
Db	634	AGCGGCCAGCGCTACACTTTCGGCAAGGTAGAGCTTC-----	666
Oy	375	GlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuLeuGlnLeuLeu	394

Db	670	-----GAGGGCGACTCGATC--ATGAGAGGAACACTGTTGGCCGCGATGGTG	714
Qy	395	ThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnSpleu	414
Db	715	CCGTTCAAGGCGCCGCAACCCCTATAGTCCGACATCATGCGCGAACTACACAGAACTG	774
Qy	415	IleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGln	434
Db	775	CAATCAACCGGCTACTTC-----	792
Qy	435	IleGlnAsnSpgInValSerPheGluGlnSerSerSerArgThrGluProIaGln	454
Db	792	-----	792
Qy	455	ValAspGluSerThrLeuGluProValIleGluThrValGluIleuThrAspGlyIleLeu	474
Db	793	-----GAGGAGTGGCG	804
Qy	475	MetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeu	494
Db	805	GTGCGAGCGCGCGGACCCAGCGTCAAGCC-----	834
Qy	495	ValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeuAlaIleAsn	514
Db	834	-----	834
Qy	515	HisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaVal	534
Db	835	-----GACGGCGCGCCGACGCGATC-----	855
Qy	535	AlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGluIaThr	554
Db	855	-----	855
Qy	555	AlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysValProLeuLys	574
Db	856	-----CCGGTACCG	864
Qy	575	ValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrPheLysAsp	594
Db	865	GTGCGGCGTGGAGGCGCGCAACCCCGGACCATGGCGCTGGCTTCGACCGAC	924
Qy	595	ThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspLysTyrGln	614
Db	925	GTGCGGCGCGCGAGGTTCAACTGACCCCGCATTTGGGTGATCCGAAAGCATAGC	984
Qy	615	AlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeuTyrAlaThrLys	634
Db	985	GTGGGATTCGAGTGGGAATCTCCGACCCAGGACAGACGTCGCGCTGATCAGATTT	1044
Qy	635	ProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGluVal	654
Db	1045	CCCTCGATCCGCGCGTACCAGAACCTCGCTTCAACGAGGCGTACCAAGTTGCAGAA--	1101
Qy	655	PheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGluHisGluIleSer	674
Db	1102	-----GACCTAGTGCACACCGAAGCAAGCATGCTGACCTTGGCGCGCA	1144
Qy	675	ArgSerIleIleGlnAsnGlyTyrPAsnArgThrTyrSerLeuArgTyrArgLeuAsp	694
Db	1147	TGGCACACAGACAGCCCGATGGCTGCGACCGGGTGGTTTCCGTCAACTGATGCGCGAG	1206
Qy	695	LysLeuLysThrGlnAlaProProGluThrTyrProLAspLeuProValAspPheValAsn	714
Db	1207	GAGTCAACAGCTG-----	1218
Qy	715	GlyLysProSer-----GlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrVal	732
Db	1219	GCGACACACTCCGGGCTGACAGCTTCATGTCGGGAGTCCGGCTACTCGCTCGAG	1278
Qy	733	AlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGly	752

DB 1279 ACCGACACAAAGGTGCGACCCAGCCATGGCTACCGCTGCAATTCACAGTTCAGAGGGGGC 1338
OY SerSerGlyLeuValSerAspAlaMetAlaIleAlaIleValGlyIleSerGlyVal 772
DB 1339 AAGGAAGGCTGTGGCCGACGCGGACGCTTCATGTGCGACGCGCAAGGGGCTG 1398
OY TyrSerPheGlyAspAlaValTyrGlySerAsnArgAlaHisGlnMetThrGlyIle 792
DB 1399 ACCAGCTTCCCGCGC-----GGCCATCGCGCTGCGGCGCGCTG 1437
OY GlnAlaIleGlyTyrTrpIleTrpSerAspAsnHisValProTyrIleArgLeuArgPhePhe 812
DB 1438 CAGGTAGGCGGAAATCGCCACCAAGCTACAGTCCGCGCTGCGCTGCGCTTCTTC 1497
OY AlAGlGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAsp 832
DB 1498 GGTGGGGGCGACAGAGCGTGGGTAGCTACCGAGCGCGCTGCGCGGAAATTCG 1557
OY LysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGlu 852
DB 1558 GATGGCGACAAAGATCGCGCGCGCTACATCATCGCGCGCTGAGTATCATATCCG 1617
OY PheMetLysAspLeuArgLeuAlaValAlaPheGlyAspIleGlyAsnAlaTyrAspLysGly 872
DB 1618 CTGCGCGAGCGCTGGCGCTGGCGACCTTGTGTGACAGGCGCAAGCCTTCAATCGCTG 1677
OY PheThrAsnAspThrLysIleGlyAlaGlyValAlaGlyValArgTyrPheAspProValGly 892
DB 1678 GACTTCGCGCTGATCAAGACCGGGGTGCGCTGCGCTGCGGTGGTCTGCGCGGTGCG 1737
OY GlnValArgValAspAlaValAlaThrGlyValLysGlnGluLysAsnProIleLysLeuHis 912
DB 1738 CCCTTGGCGCTGACCTGGCCCATGGCTGACAGACAGCGGCGGT--TTCCGCTTGAC 1794
OY 913 PhePheIleGly 916
DB 1795 TTCTCCATGGCG 1806
RESULT 3
US-10-417-886-3279 Application US/10417886
Sequence 3279, Application US/10417886
GENERAL INFORMATION:
APPLICANT: Keith G. Weinstock et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROBACTER
TITLE OF INVENTION: CLOACAE FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.135
CURRENT APPLICATION NUMBER: US/10/417,886
PRIOR FILING DATE: 2003-04-17
PRIOR APPLICATION NUMBER: US/09/252,691C
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/094,145
PRIOR FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: US 60/074,787
NUMBER OF SEQ ID NOS: 11326
SEQ ID NO 3279
LENGTH: 1860
TYPE: DNA
ORGANISM: Enterobacter cloacae
US-10-417-886-3279
Alignment Scores:
Pred. No.: 5,65e-31 Length: 1860
Score: 500.50 Matches: 183
Percent Similarity: 38.09% Conservative: 100
Best Local Similarity: 24.63% Mismatches: 253
Query Match: 10.59% Indels: 207
DB: 8 Gaps: 19
US-09-914-168-2 (1-919) x US-10-417-886-3279 (1-1860)
OY 196 ThSerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaIle 215
||||| ||| |||||

DB 169 ACCAGGGGATTCGCCAGCGC-----GCCAATGCTCCGTTTGCAG 207
OY LeuGluAspIleThr-----GlnLysSerAlaMetAsp 226
DB 208 GTTGAGGGGTATTCGGGGGCGGTGGAAAAAACGTGCGTCCGACGCTTCTACTATCCAG 267
OY LeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuVal----- 240
DB 268 AGCGATGAGGTCAAGCGCGATTCGCGCTTTCGCGCGCGGTGGATGACGCATTCGCGAA 327
OY AlaIleArgAlaValGlyTyrTyrAsp-----IleAspLeuSerIleIleArgAsnSer 258
DB 328 GCGTTAAAGCGGTGGGGGTATGATTCGCGCGCGGTGCGCGCGCGCAACCGCTGATTCGCG 447
OY IleGlyGluValAspValIleIleHisAspLeu-----GlyLysProValTyrIleAsp 276
DB 388 GCAAAAGAGCTCAGAGGTGATTCGCGCGCGGTGCGCGCGCGCAACCGCTGATTCGCG 447
OY TyrArgAlaValGluValArgGlyGlyAlaAspAspLysAlaPheThrThrValAla 236
DB 448 GGCACGACCTGTGTGCTGCGCGCGCGCGCGCGCACGCGATGCTACTGATCTGCTC 507
OY AspGluValProLeuLeuIleGlyAspValPheHisGlyLysTyrGluThrLysLys 316
DB 508 AGCAGCCGTCG---AAGTGGGACCGCTGCAATCAGCGGATTCAGATCGTTTAA 564
OY AsnLeuIleGluAsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgTyrPheAspArg 336
DB 565 AATTCATTAACAGCGTTGTGCTGGCGCAAGGCTACTTCAGACGCGCTTCAACAAAGC 624
OY SerValAspValIleLeuProAspAsnThrAlaAspValSerIleuIleTyrAspThrGly 356
DB 625 CACCTAGCGCTGCTGCTGAGCGACGACGTCATTCGAGTATCGATTTCGACAGCGGC 664
OY ThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspProLysThrAsnGlnLeu 376
DB 685 GACCGCTACCGGTTGGCGATGTGACCTTC----- 714
OY 377 ThrThrAspProAspLysLeuProValLysArgGluLeuGlnLeuLeuThrVal 336
DB 715 -----GAGGATGCGCAAAATCCGTGATGAGTATTCGCAAAACCTGGGCGCTTT 762
OY AsnMetGlyGluValTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAla 416
DB 763 AAAAAGGCGACTACTACGATCCAGCGACGACGTGGCGACGTGAACGCTTTGTCCGCG 822
OY ThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGlnIleGln 436
DB 823 ACGGCGTGTTAATCTCCGTG-----GTGGTGGCGCGGAA----- 858
OY AsnAspGlnValSerPheGlnGlnSerSerSerSerArgThrGlnProAlaGlnValAsp 456
DB 859 -----TTTGATTAATTC----- 870
OY GluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAsp 476
DB 870 ----- 870
OY IleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAla 496
DB 870 ----- 870
OY AlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAsp 516
DB 870 ----- 870
OY AspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArg 536
DB 870 ----- 870
OY AlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeu 556
DB 870 ----- 870

```

Oy 557 AlaAsnArgLysThrProAlaAspValTyrGlnSerLysLysProLeuTyrValPhe 576
Db 871 -----CGTAAACG-----AAAGTGTACCGCTGCATGCGCG 903
Oy 577 ValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrPcLysSerAspThrGly 596
Db 904 GTCTCTCCGCGCACCGAAACACCATGACCGCTGCGCTACCTCCAGCAGCGTCCGCG 963
Oy 597 ThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGly 616
Db 964 CCGCCGCTGAAGACCTCGGGAAAAAACCGTGATGTAACCTCTACGCGCCACGCTGACC 1023
Oy 617 AlaGluLeuArgLysSerGluAspLysGlyValLysLeu---TyrLathrLysPro 635
Db 1024 ACCACGCTGAGGCTGTCTCGCGCTGACACGACGCTGACTTCAGCTATTAATCGCGCTG 1083
Oy 636 LeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPhe 655
Db 1084 CTGAATAATCCCTTGACGCAATACTACTCTGTCAGGCGCGCTTAAAGCGTACC----- 1137
Oy 656 GlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHisGluLeuSerArg 675
Db 1138 -----GATTGACACGATACCGACAGACGACTCGACGCGCTT-----CGGATATCAGC 1185
Oy 676 SerIleIleGlnAsnGlyTyrPasnArgThrTyrSerLeuArgTyrArgLeuAspLys 695
Db 1186 TTCTGGGATCTCTCCAGCGGCTGGCGCGCGCATTAACCTCGCGGAGCGCTCGACAC 1245
Oy 696 LeuLysThrGlnAlaProGluThrTyrPglAspLeuProValAspPheValAsnGly 715
Db 1246 TTT---ACCCAGGCA-----AACGTC 1263
Oy 716 LysProSerGlnGluAlaLeuAlaGlyValAlaValHisLysThrValAlaAspAsn 735
Db 1264 ACTAACACACCATGCTGCTTATCCGGCGGTGATGATACGCGTACCGCGCGCGGT 1323
Oy 736 LeuValAsnProMetAlaGlyTyrArgGlnArgTyrSerLeuGlnValAlaGlySerSerGly 755
Db 1324 GGCTGTAGTCCGACCTCGGCGCGCTCTACGCGTACTCATGATATACCAACATGCGC 1383
Oy 756 LeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPhe 775
Db 1384 TGGGGCTCGACGTACTCTCCGCTTTCAGGCGCAACGTCGTGATCCGACGCTG 1443
Oy 776 GlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyIleGlnAlaGly 795
Db 1444 TAGCACAA-----CACCGCTTGTGATGCGCGGTAACTCTCGC 1482
Oy 796 TyrIleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGly 815
Db 1483 TGGATCGAAGACCGGATTTTCAGACCGCTCCGCGGATCTGGCTTTCGCGCGGGG 1542
Oy 816 AspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyr 835
Db 1543 GACCCGACGATTCGGGGGTATAGTACAATCATCTCACTGAAACACCAAGAGCCAG 1602
Oy 836 LeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAspThrGlyIuphemeLys 855
Db 1603 TTGACCGGGCGCTCAAACTGGGCGCGCTCTGCTGAGTACCACTACAGTACCGCGC 1662
Oy 856 AspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsn 875
Db 1663 AAGTGGTGGGGCGATGTTTCGTGACGGCGGTGAAGCGGTGAACATATCCGCCGACG 1722
Oy 876 AspThrLysIleGlyAlaGlyValGlyValArgTyrPalaSerProValGlyGlnValArg 895
Db 1723 GATTTCAAAACCGCGCGCGGTAGCGCTGCGTACGACGCGGTGCGGCCCAACAG 1782
Oy 896 ValAsp-----ValAlaThrGlyValLysGluGluGlyAsnProIleLysLeuHisSpe 913
Db 1783 CTCGATTTTCCGCTTCCGCTGGGCGACAAAGCAACAGCGGT-----TTACAGTTT 1833

```

```

Oy 914 PheIleGly 916
Db 1834 TACATCGCT 1842

RESULT 4
US-09-913-101-1
; Sequence 1, Application US/09913101
; GENERAL INFORMATION:
; APPLICANT: Ruelle, Jean-Louise
; APPLICANT: Thonnard, Joelle
; TITLE OF INVENTION: Haemophilus Influenzae RD Outer Membrane
; FILE REFERENCE: BM45368
; CURRENT APPLICATION NUMBER: US/09/913,101
; PRIORITY FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: 9902880.5
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: PCT/EP00/00887
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
US-09-913-101-1

Alignment Scores:
Pred. No.: 8,87e-26 Length: 1737
Score: 436,50 Matches: 142
Percent Similarity: 35,64% Conservative: 105
Best Local Similarity: 20,49% Mismatches: 235
Query Match: 9,23% Indels: 211
DB: Gaps: 14

US-09-914-168-2 (1-919) x US-09-913-101-1 (1-1737)
Oy 243 ArgAlaValAlaGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer---IleGlyGlu 261
Db 217 CGGTGTTTGGTTATATATGATCTTCCGCGCTTTGAAACGAAACAGCGTCAAGC 276
Oy 262 ValAspValIleIleHisAspLeu-----GlyGluProValTyrIleAspTyrArgAla 279
Db 277 CGCGATTTATGATGCTCATGCTATGACCCAGCAGCCAAACAAATTCGCGGAGCTGAT 336
Oy 280 ValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThrValAlaAspGluVal 299
Db 337 GTCCAAATTGAGGGGAGACCCGCAAGATGAATTTTAATGCAATTCGCTAAACCTTA 396
Oy 300 ProLeuLeuIleGlyAspValPheHisGlyLysGlyTyrGluThrLysLysAsnLeuIle 319
Db 397 CCG---AAGATGCGCTTTTGGTGGACCAACCAACTACGATGATTAACAACAGCGATT 453
Oy 320 GluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPleuAspArgSerValAsp 339
Db 454 TCACGTTTAGCATTAATGCTGGGTATTTTGATGGGAGACTTAAATTTTCACGTTAGAA 513
Oy 340 ValIleLeuProAspAsnThrAlaAspValSerLeuIleGlyTyrAspThrGlyThrGlnTyr 359
Db 514 ATGAGCCCTGAACCCATCAAGCATGCTGGCGGATGTTATTTGATAGTGCTCGCTTAT 573
Oy 360 ArgPheAspGluValAlaPhePheThrIleAspProLysThrAsnGlnLeuThrThrAsp 379
Db 574 CATTAATGCAATATTAATTT----- 594
Oy 380 ProAspLysLeuProValLysArgGluLeuLeuGlnGlnLeuLeuThrValAsnMetGly 399
Db 595 ---AGCCATTCAACAATTCGGGATGATTAATTAATTAATTCCTTAACAATCAATCTG 651
Oy 400 GluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAspLeuLeuAlaThrArgTyr 419
Db 652 GATCGTATTTAATTAATTAATTTGTCGATTTTAACGACGATTTTCCATCTCAAAATTGG 711

```

Oy	420	pheInsmetValaInshrcIuLeValaPheProGluAcrgGluGlnIleGlnAsnAspGln	439
		:	
Db	712	TTTATGCGCACTA-----	723
Oy	440	ValSerPheGluInSerSerSerSerArgThrGluProAlaGlnValaAspGluSerThr	459
Db	723	-----	723
Oy	460	LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspLleSerPro	479
Db	723	-----	723
Oy	480	IleGluPheSerAlaSerAsnLeuIleGlnAspLysLleAsnLeuValAlaAlaLysAla	499
		:	
Db	724	-----TTAGTTCAGCCTAATGTTAAI-----	744
Oy	500	ArgHisLeuThrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyVal	519
Db	744	-----	744
Oy	520	AsnArgSerIleLeuGluArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeu	539
Db	744	-----	744
Oy	540	ProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArg	559
Db	744	-----	744
Oy	560	LysThrProAlaAspValTYrGlnSerLysLysValProLeuTYrValPheValAlaSer	579
		: :	
Db	745	-----CATAAACCAAAACCTGGATGGAGCAATATTTCTTTATCA	786
Oy	580	AspLysProAlaAspGlyGlnIleGlyLeuGlyTYrPGLysSerAspThrGlyThrArgLeu	599
		: :	
Db	787	CGTAAACCAAAATGCGATGGAGCACTGGTGCGCTTTTCTACTGATGGCGCGCTTACGGA	846
Oy	600	ValThrLysPheGluHisAsnLeuIleAsnAlaAspGlyTYrGlnAlaGluLeu	619
		: :	
Db	847	CAATATAGCGCTGCAACAAACCTGGATTAATATAGCCGTGGACATAGTTGGCTTCAATCTT	906
Oy	620	ArgLeuSerGluAspLysLysGlyValLys--LeuTYrAlaIleThrLysProLeuSerHis	638
		: :	
Db	907	TATCTCTCTGCAACCAAAACCAACTGTAGAGCAACTATTCAGTAATGCCACTGCTTAAATAT	966
Oy	639	ProLeuAsnAspGlnLeuAlaArgAlaThrLeuGlyTYrGlnGlnGluValPheGlyHisSer	658
		: :	
Db	967	CCATTAAATTTATCTAGTGAATTTGGCGCTGCGTGGGA-----GGCGAAAA	1014
Oy	659	ThrAsnGlyPheAspLeuSerThrArgTYrThrLeuGlnHisGluLleSerArgSerIleIle	678
		:	
Db	1015	GAGCAT-----GATACCCAATACGAGAGGCTCTTACGTCACGCTTACGTAATGGAAT	1066
Oy	679	GlnAsnGlyGlyTYrPAsnAlaGlyThrTYrSerLeuAlaGTYrArgLeuAspLysLysLeuTYrThr	698
		: :	
Db	1069	AATGCGCATCGTTGGCAATATTTTGGCGGACCTTCGATGGAGTACGACAGCTTT--ACA	1122
Oy	699	GlnAlaProProGluThrThrPGLAspLeuProValaAspPheValaAsnGlyLysProSer	718
		: :	
Db	1126	CAAGCGGAT-----	1133
Oy	719	GlnGluAlaLeuLeuAlaGlyValaIaIaValaHisLysThrValAlaAspAsnLeuValaSn	738
		: :	
Db	1135	-----ATCAGTCGATAAACCTTA-----CTTCTTTAT	1166
Oy	739	ProMetArgGlyTYrArg-----	744
		:	
Db	1162	CCAACCTGTTGGATTACTCGTACTCGCATTCAGTGGTGCTCTTGGCCACTGGGGCGAT	1221
Oy	745	---GlnArgTYrSerLeuGluValaGlySerSerGlyLeuValSerAspAlaAsnMetAla	763
		: : :	
Db	1222	GTCGCAAAATAATTAATCTTTGATTTAAGCAAAAGCAATTTGGCTATCAGCAATCTCTTTATA	1281
Oy	764	IleAlaArgAlaGlyIleSerGlyValTYrSerPheGlyAspAsnAlaTYrGlySerAsn	783

```

Db 1282 AATGTCACACATCTACGCGCGGTGCTGACTATTGACGAAAT----- 1326
Oy 784 ArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrSerAspAsnPhasn 803
Db 1327 -----CATGCTGTCGTGCTGCTGCTGAAATCGGGTATTTTCATACAAACGTATTGAA 1380
Oy 804 HisValProGlyArgLeuAlaGlyPheAspPheAlaGlyIleGlyAspGlnSerIleArgIleTyrAla 823
Db 1381 AAATTTCCGCGCTACACGCGCTTCTTTGCGTGGGAGCATGTAAGTGTGCGCGGTACCGC 1440
Oy 824 HisAspSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlyIleValLeuAla 843
Db 1441 TATATAAAAAATTTGGCGCTATAAAATACAAATATGGAAATTTGGTGGGTTCGCGCAATTCG 1500
Oy 844 ValIleThrAlaGlyGlyTyrAsnTyrGlyIlePheMetLysAspLeuArgLeuAlaIlePheGly 863
Db 1501 ACCACTCTTTAGATATATCAATATCAAGTGTATCCAAATTTGGTGGCGCGCACTTTTGA 1560
Oy 864 AspIleGlyAsnAlaTyrAspLysGlyIlePheThrAsnAspThrLysIleGlyAlaGlyAla 883
Db 1561 GATAGTGGATAGCTGCGCATTAATTCACAGCAAAAGAGCTGGATTAGGCACAGCGCTT 1620
Oy 884 GlyValArgTyrPalaSerProValGlyIleValArgValAspAlaIleThrGlyValLys 903
Db 1621 GGTGGCGGTGGCGCATGCCACAGTTGGCGGATTAATTTGATATTTGCGCACACCCATTGCT 1680
Oy 904 GluGluGlyAsnProIleLysLysIleHisPhePheIleGly 916
Db 1681 GATTAAGTAAACACGCAAAATATTCAATTATTACATGCGA 1719

RESULT 5
; Sequence 3, Application US/09913101
; GENERAL INFORMATION:
; APPLICANT: Ruelle, Jean-Louise
; APPLICANT: Thonnard, Joelle
; TITLE OF INVENTION: Haemophilus Influenzae RD Outer Membrane
; FILE REFERENCE: BM45368
; CURRENT APPLICATION NUMBER: US/09/913.101
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: 9902880.5
; PRIOR FILING DATE: 1999-02-09
; PRIOR APPLICATION NUMBER: PCT/EP00/00887
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1731
; TYPE: DNA
; ORGANISM: Haemophilus Influenzae
US-09-913-101-3

Alignment Scores:
Pred. No.: 4,81e-25 Length: 1731
Score: 427.50 Matches: 141
Percent Similarity: 35.64% Conservative: 106
Best Local Similarity: 20.35% Mismatches: 235
Query Match: 9.04% Indels: 211
DB: Gaps: 14

US-09-914-168-2 (1-919) x US-09-913-101-3 (1-1731)
Oy 243 ArgAlaValAlaGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSer---IleGlyLu 261
Db 211 CGTGTGTTGGTTAATTAAGAATCTTCGCTGCTTTCGACCAAAACGCGTCAAGGCATAA 270
Oy 262 ValAspValIleIleHisAspLeu-----GlyGluProValTyrIleAspTyrArgAla 279
Db 271 CGCGATTTATTATTTGCTCATCTTACACACAGCGACGCAACAAATTTGGGGGACATGAT 330
Oy 280 ValGluValArgGlyGluGlyAlaAspLysAlaPheThrThrValAlaAspGluVal 299

```



```

Db 331 GTGCAATGAGGGGACCCGACAAAGATGAAATTTTGATGCGCTACGTAATAACTTG 390
OY 300 ProLeuLeuIleGlyAspValPheHisIleGlyLysTyrGluThrLysLysAsnLeuIle 319
Db 391 CCA---AAGAAGAGCGCTTGTGTGAACACCAAACTTACGATCATTAACAAACGCGATT 447
OY 320 GluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrLeuAspArgSerValAsp 339
Db 448 TCACCGCTGGCATTAATGCTGGTATTTTGATGGGAACTTTAAATTTACCTTTAGAA 507
OY 340 ValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyr 359
Db 508 ATCAGCCCTGAAACCCATCAACGATGTCGCGCATTTGATTTGATGATGTCGTCGCTAT 567
OY 360 ArgPheAspGluValAlaPhePheThrIleAspProLysThrAsnGlnLeuThrThrAsp 379
Db 568 CATTATGGCAATATTACTTT----- 588
OY 380 ProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGly 399
Db 589 ---AGCATTCACAAATCGGTGACGATTCATCAATATATCTTAACTCAATCAATCTGGC 645
OY 400 GluAlaTyrAsnLeuGlnAlaValAlaGlaLeuSerAsnAspLeuIleAlaThrArgTyr 419
Db 646 GATCCATATTTATGATGATATTTGCGGATTAACCGACGATTTTTCATCTCAAAATTTGG 705
OY 420 PheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGln 439
Db 706 TTATAGCTCAGTA----- 717
OY 440 ValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThr 459
Db 717 ----- 717
OY 460 LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspLysPro 479
Db 717 ----- 717
OY 480 IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAla 499
Db 718 -----TTAGTTCAGCCCAATGTTAAT----- 738
OY 500 ArgHisLeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyVal 519
Db 738 ----- 738
OY 520 AsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeu 539
Db 738 ----- 738
OY 540 ProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArg 559
Db 738 ----- 738
OY 560 LysThrProAlaAspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSer 579
Db 739 -----CATAAAGCAAAACCTGGATGCGAGATTTTCTTTTCA 780
OY 580 AspLysProArgAspGlyGlnIleGlyLeuGlyTyrPheGlySerAspThrGlyThrArgLeu 599
Db 791 CGTAAATAAATAATGCGATGGAACCTGGCTGGCTTGTGCTACTGATGCGCGCTTCACGGA 840
OY 600 ValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaIleGlyAlaGluLeu 619
Db 841 CAAATTTGGCTGGACAAACCTTGATTAATAGCCGGGACATAGTTTGGCTTCAAAATCTT 900
OY 620 ArgLeuSerGluAspLysLysGlyValLys---LeuTyrAlaThrLysProLeuSerHis 638
Db 901 TATCTCTGTCGACCAAAACACTAGAGGCAACTTATCGAATGCGCACTGCTTAAAT 960
OY 639 ProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValAlaPheGlyHisSer 658

```

```

Db 961 CCATTAATTAATTAATGATTTTCCGTCGCTGGGAA-----GGGCAAAA 1008
OY 659 ThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIle 678
Db 1009 GAGAAC-----GATACCATACGAGAGTGTACCTTGTACCGCTTATTTGGAAT 1062
OY 679 GlnAsnGlyTyrPheAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThr 698
Db 1063 AATGCCCATGTTGGCAATATTTTGGCGGACTTCGTCACGCGATACACACACTTT---ACA 1119
OY 699 GlnAlaProGluThrThrPheGlnAspLeuProValAspPheValAsnGlyLysProSer 718
Db 1120 CAAGCGAT----- 1128
OY 719 GlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuValAsn 738
Db 1129 -----ATCACTGATTAACCTTA-----CTTCTTTAT 1155
OY 739 PrometArgGlyTyrArg----- 744
Db 1156 CCAACTGTTGATTTACTGCGCACTGATTAAGTGTGTTCTTCCACTTGGCGCAT 1215
OY 745 ---GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAla 763
Db 1216 GTGCAAAAATTAATTAATTTATTAACCAACGAATTTGGCTATACAGAACTTCTTTATA 1275
OY 764 IleAlaArgAlaGlyLysSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsn 783
Db 1276 AAGTCAAGACATCTAGCCGCGGTGCTACTTATGCAGAAT----- 1320
OY 784 ArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyrIleLysPheSerAspAsnPheAsn 803
Db 1321 -----CATCGTATCTGCTGCTGCTGAATGCGGATTTTACATCAAAAGATATTGAA 1374
OY 804 HisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAla 823
Db 1375 AAAATTCGCCCAACACTGCTGCTGCTGCTGCGGATCTGCGCGGCTTTCAGGC 1434
OY 824 HisAspSerLeuSerProLysSerAspLysGlyTyrLeuThrGlyGlyGlnValAla 843
Db 1435 TATAAAAAATTCGCGCTTAAATTAAGAAATGCAAAATTTGGCTGCGGCTCAGCTTGCCT 1494
OY 844 ValGlyThrAlaGlyTyrAsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGly 863
Db 1495 ACCACTTCTTTAGAAATCAATATCAAGTTTATCCAAATTTGCGGCGCACTTTTGA 1554
OY 864 AspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGlyVal 883
Db 1555 GATAGTGGATTAGCTGCCGATTAATTCACAGCAAAAGAGCTGCTTATGGCACAGCGCTT 1614
OY 884 GlyValArgThrAlaSerProValGlyGlnValArgValAspValAlaThrGlyValLys 903
Db 1615 GGTGTCGTTGGCATCGCCAGTGGGTGGCATTTAAATTTGATATTCGCCACACCATTCGT 1674
OY 904 GluGluGlyAsnProLysLeuHisPhePheIleGly 916
Db 1675 GATAAAGATTAACGCAAAATATTCAAATTTTACATCGGA 1713

RESULT 6
US-10-419-128-14663
; Sequence 14663, Application US/10419128
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/10/419,128
; PRIOR FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US/09/252,991
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

```



```

Oy 774 SerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGln 793
    |||||
Db 419 ACCTTCGGCGGC-----GGCCATCGGCTGCTGCGCGCGCTGCAG 457
Oy 794 AlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAla 813
    |||||
Db 458 GTAGCGGAGATGCGCAGCAGCACTACAGTGCATCCGCCCTCGCTGCGCTTCCTGCT 517
Oy 814 GtlyGlyAsnGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLys 833
    |||||
Db 518 GCGCGGCGACACAGCGCTGCTGCTGCTACGACCTGCTGCTGCTGCTGCTGCTGCTGCT 577
Oy 834 GtlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGlyPhe 853
    |||||
Db 578 GCGCACAAGATGCGCGCGCTACATGATCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 637
Oy 854 MetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPhe 873
    |||||
Db 638 GCGGAGCGCTGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 697
Oy 874 ThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTyrAlaSerProValGlyGln 893
    |||||
Db 698 TTCGCGTGCATGACAGACCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 757
Oy 894 ValArgValAspValAlaThrGlyValLysGlnGlyLysAsnProIleLysLeuHisPhe 913
    |||||
Db 758 TTGCGCGCTGACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 814
Oy 914 PheIleGly 916
    |||||
Db 815 TCACATGCGG 823

```

```

RESULT 8
US-10-419-128-14966/c
; Sequence 14966, Application US/10419128
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/10/419,128
; PRIOR FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US/09/252,991
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14966
; LENGTH: 726
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-419-128-14966

```

```

Alignment Scores:
Pred. No.: 3,566-20 Length: 726
Score: 362.00 Matches: 82
Percent Similarity: 51.67% Conservative: 42
Best Local Similarity: 34.17% Mismatches: 90
Query Match: 7.66% Indels: 26
DB: 8 Gaps: 4

```

```

US-09-914-168-2 (1-919) x US-10-419-128-14966 (1-726)
Oy 679 GlnAsnGlyGlyTyrPasnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuTyrThr 698
    |||||
Db 672 CAGCCCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 613
Oy 699 GlnAlaProGluThrTyrPasnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuTyrThr 718
    |||||
Db 612 -----GGCCAGCAGCTCC 601

```

```

Oy 719 -----GlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeu 736
    |||||
Db 600 GGGCTAGACAGCTTCTGTATGCCGGGATCGGCTACTCCCTGCTGGAGACCAACAAAG 541
Oy 737 ValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGlnValGlySerSerGlyLeu 756
    |||||
Db 540 GTGACCCCGACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481
Oy 757 ValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValIleSerPheGly 776
    |||||
Db 480 CTGGCCGACCGACGCTCTCCATGCTGACAGCGCATGCCAAGGCGCTGACCACTTCCG 421
Oy 777 AspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyr 796
    |||||
Db 420 GGC-----GGCCATGCGCTGCTGCGCGCTGCGCGCTGCGCGCTGCGCGCTGCG 382
Oy 797 IleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyIleAsp 816
    |||||
Db 381 ATCGCCACCAAGCACTACAGTGCATCCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 322
Oy 817 GlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeu 836
    |||||
Db 321 CAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 262
Oy 837 ThrGlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGlnLysPheMetLysAsp 856
    |||||
Db 261 ATCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 202
Oy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAsp 876
    |||||
Db 201 TCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 142
Oy 877 ThrLysIleGlyAlaGlyValArgTyrAlaSerProValGlyGlnValArgVal 896
    |||||
Db 141 ATCAACACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 82
Oy 897 AspValAlaThrGlyValLysGlnGlyLysAsnProIleLysLeuHisPhePheIleGly 916
    |||||
Db 81 GACCTGCGCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 25

```

```

RESULT 9
US-10-366-683-14966/c
; Sequence 14966, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubenfield, Marc J.
; APPLICANT: Nollinger, Jork
; APPLICANT: Delouhery, Craig
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: PAT#03-04
; CURRENT APPLICATION NUMBER: US/10/366,683
; PRIOR FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14966
; LENGTH: 726
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-366-683-14966

```

```

Alignment Scores:
Pred. No.: 3,566-20 Length: 726
Score: 362.00 Matches: 82
Percent Similarity: 51.67% Conservative: 42
Best Local Similarity: 34.17% Mismatches: 90
Query Match: 7.66% Indels: 26
DB: 9 Gaps: 4

```

```

US-09-914-168-2 (1-919) x US-10-366-683-14966 (1-726)
Oy 679 GlnAsnGlyGlyTyrPasnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuTyrThr 698

```

```

Db 672 CAGCCGATGCTGCGACGCGGTGTTTCGCTGAATGATGCCGAGAGTACAACTG 613
Oy 699 GlnAlaProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysProSer 718
Db 612 -----GlnGluAlaLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeu 736
Oy 719 -----GlnGluAlaLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeu 736
Db 600 GGGCTGAGCAGCTTCCTGATGTCGGGGATCGGCTACTCGCTCGGAGACCGACAAACAG 541
Oy 737 ValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySerSerGlyLeu 756
Db 540 GTGACCCCAAGCCATGGCTACCGGCTGACAGTCAACGTCAAGGGGGGAGGAAGGAGCGCTG 481
Oy 757 ValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGly 776
Db 480 CTGGCCGACGCGGACGCTCCATGTCGACGCCCAAGGCGCTGACCAAGCTTGCC 421
Oy 777 AspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAlaGlyTyr 796
Db 420 GGC-----GGCCATCGCTGCTCGGCGCGCTGACAGTAGGCGGA 382
Oy 797 IleTrpSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyLysAsp 816
Db 381 ATCCGCAACAGCACTACAGTCCGCTCGGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 322
Oy 817 GlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeu 836
Db 321 CAGACCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 262
Oy 837 ThrGlyGlnValLeuAlaValAlaGlyThrAlaGluTyrAsnTyrGluPheMetLysAsp 856
Db 261 ATCGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 202
Oy 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThrAsnAsp 876
Db 201 TCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 142
Oy 877 ThrLysIleGlyAlaGlyValAlaGlyValArgTyrPheAsnProValGlyGlnValArgVal 896
Db 141 ATCAAGACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 82
Oy 897 AspValAlaThrGlyValLysGluGluGlyAsnProIleLysLeuHisAspPheIleGly 916
Db 81 GACCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 25

RESULT 10
US-10-144-771-21979
: Sequence 21979, Application US/10144771
: GENERAL INFORMATION:
: APPLICANT: VENTNER, J. Craig
: TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
: FILE REFERENCE: CL001321
: CURRENT APPLICATION NUMBER: US/10/144,771
: CURRENT FILING DATE: 2002-05-15
: NUMBER OF SEQ ID NOS: 47235
: SEQ ID NO 21979
: LENGTH: 1416
: TYPE: DNA
: ORGANISM: HUMAN
US-10-144-771-21979

Alignment Scores:
Pred. No.: 4,6e-19 Length: 1416
Score: 353.00 Matches: 148
Percent Similarity: 32.63% Conservative: 85
Best Local Similarity: 20.73% Mismatches: 215
Query Match: 7.47% Indels: 266
DB: Gaps: 16

US-09-914-168-2 (1-919) x US-10-144-771-21979 (1-1416)

```

```

Oy 211 AsnIleLysAlaAlaLeuGluAspIleThrGlnGluSerAlaMetAspLeuAsnGlySer 230
Db 31 AACGTCGTGGCGACGCTTTCTACGATTGAAAGTATGAAAGGACCGACACCGCGCTTT 90
Oy 231 IleProArgLeuArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyr----- 248
Db 91 CGCGCAGCGCTGATGATGTCATCCGCGGAGGTCGTAAGCGCGGCTTTTACACCGCG 150
Oy 249 -----AspIleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIle 266
Db 151 ACCATTGAATTTGATTCCTCCGTCACCGCCCAAGAAAGCGCGCGAGATTGATGCCAAA 210
Oy 267 HisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGly 286
Db 211 GTCACGCGACGCGCGGCTGTTAATGGCGGACCGCATGTTATTCGGCGCGCGCG 270
Oy 287 AlaAspLysAlaPheThrThrValAlaAspGluValProLeuIleGlyAspVal 306
Db 271 CGGACCGATAAAGACTATTGAAATGCTCGATACTCGCCG--GCTATTGGCAGCGTA 327
Oy 307 PheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHis 326
Db 328 CTGAACCAAGCGCGATTTATGAAATTTCAAAAGTCCTTAACCAACATTCGCTGCTAAA 387
Oy 327 GlyTyrPheAspGlyArgTyrPheLysAspArgSerValAspValIleLeuProAspAsnThr 346
Db 388 GGTATTTCGATACGATTTACCAAGCGCAGCTGGGCAATTCGCGCTGCTGCTGCTGCTGCTG 447
Oy 347 AlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhe 366
Db 448 GCCTTCGGCATATTGATTATTAACAGTGGCGAACGTTTACCGCTTTGGGACGTACCTTT 507
Oy 367 PheThrIleAspProLysThrAsnGlnLeuThrThrAspAspLysLeuProValLys 386
Db 508 -----GAAGATCAACAATCCGC 525
Oy 387 ArgGluLeuLeuGluGlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAla 406
Db 526 GATGAATACTCGCAAAATCTGCTGCTGCTTAAAGAGGCGGTGATGCAATTCGAAGAT 585
Oy 407 ValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGlu 426
Db 586 CTGGCAGACTGAAACCGCGCTTCTGCTACCGCGCTGCTTAATAG-----#----- 633
Oy 427 IleValPheProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSer 446
Db 633 ----- 633
Oy 447 SerSerArgThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThr 466
Db 634 -----CCGCTGCTGAG-- 645
Oy 467 ValGluLeuThrAspGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsn 486
Db 646 -----AATCCACGCGAA-----CAATTATTAT 666
Oy 487 LeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetPro 506
Db 667 TTGCTGCAG----- 675
Oy 507 AspAspArgValLeuAlaIleAsnHisAspArgValaAsnArgSerIleLeuGlyArg 526
Db 676 -----GGCGGTTTAAAGCGACT----- 693
Oy 527 IleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSerGluAsnGlu 546
Db 694 -----GACCTGAACGATACCGAA 711
Oy 547 ValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyr 566
Db 711 ----- 711
Oy 567 GlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArgAspGlyGln 586

```

```

Db 711 ----- 711
Oy 587 IIEGLYLEUGLYTPRGLYSERASPRHGRGLTHRARGLEUVALTHRLYSPHEGLUNHISA 606
Db 712 ----- 744
Oy 607 LEUILEASARGSPGLYTRGILNAGLVALLEUARGLEUSERGLUSPGLYS 626
Db 744 ----- 744
Oy 627 GLYVALYSPLEUTRALATHRLYSPROLEUSERHISPROLEUASNPGLINLEUARGALA 646
Db 744 ----- 744
Oy 647 THILEUGLYTRGILNGINGLVALPHEGLYHISSETHIASNGLYPHEASPLEUSERTHR 666
Db 745 ----- 759
Oy 667 ARGTHLEUGLUNHISGLUILESERARGSERILEGLIASNGLYLYTPRASNARGTHR 686
Db 760 ----- 774
Oy 687 TYRSERLAUGTYRARGLEUASPLYSLEULYTHRGILALAPROGLUTHRTRGILN 706
Db 775 ATTAACSTGCGCTGGAGTCTGCAC----- 798
Oy 707 ASPLAUPROVALASPRHEVALASNGLYLVSPROSERGLINGLUALALEU-----Ala 724
Db 799 ----- 846
Oy 725 GLYVALALAVALHISLYSTHRVALALASPAISLEUVALASNPROMETARGLYTYRARG 744
Db 847 GGGGATGATATAGCCGACGCGCTCTCGTGCGCCATGACCAACSTGGGCGACSTGC 906
Oy 745 GILNAGTYRSERLAUGLYVALGELYSERGLYLEUVALSERASPLASMETALALIE 764
Db 907 CAAACSTACTATGACACTCTCCAAACAGGCGCTGAGATGTCATTTCTCCGTT 966
Oy 765 ALAARGALAGLYLESERGLYALTYRSERPHEGLYASPAHALATRGLYSERASARG 784
Db 967 TTCACGCGCCACAGCTCGATCCGACACTGATACGAC----- 1008
Oy 785 ALAHISGLIMETHTHRLGLYLTIEGLNALAGLYTYRILETRSERASPAISPHASNHIS 804
Db 1009 ---CATTCGTTTGTTACACGCGGACGCTGGGTGATGAACCGGTATTTCCACAA 1065
Oy 805 VALPRTGYRARGLEUARGPHERHEALAGLYGLYASPLISERTILEARGLYTYRALAHIS 824
Db 1066 GTACCGCGGATCTGCGTTCTTCGCGGCGGACCGCAGATTCGTGCTACAAATAC 1125
Oy 825 ASPSERLEUSERPROLESERASPLYSGLYTYRLEUTHRGLYGLYINVALLEUHALVAL 844
Db 1126 AANTCTATGCTCCGAATATACCCCAACGGTGACCTGAAGAGGCGCTGAGATTATACC 1185
Oy 845 GLYTHRALAGLYTYRASNTRYRGLUPHEMELYSASPLEUARGLEUVALPHEGLYASR 864
Db 1186 GGATGCGTCGATACCAAGACGACGACGGAATGCGGCGCGCGGTGTTTCGAT 1245
Oy 865 IIEGLYASNALATYRASPLYSGLYPHERTHASNPSTHLYSILEGLYALAGLYVALGILY 884
Db 1246 AGTGGCAAGCGGTAAAGCATATTCGCGCAGCAGCATTTAAACCGCTACCGGCGGCG 1305
Oy 885 VALARGTRALASERPROVALGILYINVALARGVALASPRVALALATHRGLYVAL----- 902
Db 1306 GTGCTCGGGAATCGCGGTGCGGCAATCAATCAATGATTTTCCCTACCGCGTGGGAT 1365
Oy 903 LYSGLINGLYLASNPROLLEUSERLEUHSIPHEHILEGLY 916
Db 1366 AAAGACGAACACGGG-----TTACAGTTTACATCGCT 1398

```

RESULT 11
US-60-453-134-81

```

; Sequence 81, Application US/60453134
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF H. INFLUENZAE
; FILE REFERENCE: 28335/38815
; CURRENT APPLICATION NUMBER: US/60/453,134
; CURRENT FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 588
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 81
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: H. Influenzae
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1309)..(1309)
; OTHER INFORMATION: n = a, c, g, or t
US-60-453-134-81

```

```

Alignment Scores:
Pred. No.: 6,39e-19 Length: 1365
Score: 351.00 Matches: 98
Percent Similarity: 43.87% Conservative: 63
Best Local Similarity: 26.70% Mismatches: 148
Query Match: 7.43% Indels: 58
Gaps: 8
DB: 10

US-09-914-168-2 (1-919) x US-60-453-134-81 (1-1365)
Oy 566 TYRGLSERLYSLYSVALPROLEUTYRVALPHEVALALSERASPLYSPROARGSPGLY 585
Db 79 CATAAAGTAAACCGGTATAGTAATCATTTATTCACGTAATAAATCGCATG 138
Oy 586 GILIEGLYLEUGLYTPRGLYSERASPRHGRGLTHRARGLEUVALTHRLYSPHEGLUNHIS 605
Db 139 GAACCTGEGTGCGCTTCTACTGATGGCGGCGTTACACGACAAATAGCGTGCACAA 198
Oy 606 ASNLEUILEASARGSPGLYTYRGLNALAGLYALGILEUARGLEUSERGLUSPGLYS 625
Db 199 CCGTGATTAATAGCGCGTGACATAGTTGCGTTCAATCTTATCTCTGACACAA 258
Oy 626 LYSGLYVALYS-----LEUTYRALATHRLYSPROLEUSERHISPROLEUASNPGLINLEU 644
Db 259 CAACACTGTAGAGCAACTTATCGAATCCACTCTTAAATCCATTAATATTACTAT 318
Oy 645 ARGALATHRLLEUGLYTYRGLINGLVALPHEGLYHISSETHIASNPGLINLEUASPL 664
Db 319 GATTTTCCGTCGCTGGGAA-----GGGGAACAGAAC-----GATACC 360
Oy 665 SEETHARGTHRLLEUGLUNHISGLUILESERARGSERILEIEGLIASNGLYTYRASN 684
Db 361 AATACGACAGTGTTCACGTTTCACGCGTTACGTTATGCAATATACGATGTTGGCA 420
Oy 685 ARGTHTYRSERLAUGTYRARGLEUASPLYSLEULYTHRGILALAPROGLUTHR 704
Db 421 TATTTGCGGACTTCGTACGCGATACGACAGTTT---ACACAAACGGAT----- 468
Oy 705 TRPGLNASPLEUPROVALASPRHEVALASNGLYLVSPROSERGLINGLUALALEU 724
Db 468 ----- 468
Oy 725 GLYVALALAVALHISLYSTHRVALALASPAISLEUVALASNPROMETARGLYTYRARG 744
Db 469 -----ATCAGTGAATAAACCTTA-----CTTCTTATTCCAACTGTGGATTACT 513
Oy 745 ----- 749
Db 514 CGTACTGATTAAGTGTGTTCCCTTGTGCCACTTGGGCGAGTGTCAAAAAATTACTTT 573
Oy 750 GLUVALGILYSERSERLYLEUVALSERASPLASMETALALIEALARGALAGLYLIE 769
Db 574 GATTTAAGCAAAAGATTGGGTATCAGATCTTCTTTATTAACAGTGAACGATCTAC 633

```

```

QY 770 SerGIyAlaIyTyrSerPheGlyAspAsnAlaIyGlySerAsnArgAlaHisGlyMetThr 789
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 634 GCGTCGGTTCCTACTTATGCGAAAT-----CATGATATCGTT 672
QY 790 GLyGlyIleGlnAlaGlyTyrIleTyrSerAsnPheAsnHisValProTyrArgIleu 809
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 673 CCGTCGTCGCAAAATCGGTATTTACATACAAAGCATATTGAAAAATTCGCCCTACACTG 732
QY 810 ArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerPro 829
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 733 CGTTTCTTCTGCGGCGGATCGTAGTCGCGGTTACGGCATATAAAAAATTCGCCCT 792
QY 830 IleSerAspLyGlyTyrLeuThrGlyGlyGlnValIleuAlaValGlyThrAlaGlyIyr 849
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 793 AAAAATAAAACGCGAAATTAAGTTCGTCGCGATTCGCTTACCGCTTCTTAAAGAT 852
QY 850 AsnTyrGluPheMetLysAspLeuArgLeuAlaValIlePheGlyAspIleGlyAsnAlaTyr 869
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 853 CAATATCAAGCTTATGCAAAATTCGTCGCGGCACTTTTGCAGATAGTGATAGTCTGCC 912
QY 870 AspLyGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTyrAlaSer 889
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 913 GATTAATTCACGCGCAAAAGATTCGTATGTCGAGCGGTGTGTGTCGTCGCGCATCG 972
QY 890 ProValGlyGlnValArgValAspValAlaThrGlyValIleGlyGluGlyAsnProIle 909
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 973 CCAGTGTGCGCATTAATTTGATATGTCACACACCCATTCGTGATTAAGATACAGCAAA 1032
QY 910 LysLeuHisPhePheIleGly 916
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 1033 AATATTCATTTTACATTCGA 1053

```

```

RESULT 12
US-10-399-416-9
; Sequence 9, Application US/10399416
; GENERAL INFORMATION:
; APPLICANT: Thonard, Joelle
; TITLE OF INVENTION: BASB207 Polypeptides and Polynucleotides
; TITLE OF INVENTION: from NonTypeable Haemophilus Influenzae
; FILE REFERENCE: BM45424
; CURRENT APPLICATION NUMBER: US/10/399,416
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: PCT/EP01/11983
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: GB 0025488.8
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 944
; TYPE: DNA
; ORGANISM: non-typeable Haemophilus influenzae
US-10-399-416-9

```

```

Alignment Scores:
Pred. No.: 1,64e-16 Length: 944
Score: 319.00 Matches: 93
Percent Similarity: 44.38% Conservative: 61
Best Local Similarity: 26.80% Mismatches: 135
Query Match: 6.75% Indels: 59
DB: 8 Gaps: 8

```

US-09-914-168-2 (1-919) x US-10-399-416-9 (1-944)

```

QY 586 GlnIleGlyLeuGlyTyrPGLysAspThrGlyThrArgLeuValIleThrLysPheGlnHis 605
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 3 GAATCGGTGGCTTCTGCTACTGATGGCGGCGCTTACGAGCAAAATAGGTGACAAAA 62
QY 606 AsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuAlaGlyLeuSerGluAspLys 625
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 63 CGTTGATTAATACCGCTGACATAGTTTGGCTTCAATCTTTATCTCTGCGACACAAA 122
QY 626 LysGlyValLys---LeuTyrAlaIleThrLysProLeuSerHisProLeuAsnAspGlnLeu 644

```

```

Db 123 CAATCTAGAGGCACTTATGCAATGCGACTGTTAAATATCCATTAATATTACTAT 182
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 645 ArgAlaThrLeuGlyTyrGlnGlnGlyValPheGlyHisSerThrAsnGlyPheAspLeu 664
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 183 GATTTTGGCTCGGTTGGAA-----GGGGAAGAGAAC-----GATATCC 224
QY 665 SerThrArgThrLeuGlnHisGlnIleSerArgSerIleIleGlnAsnGlyTyrPheAsn 684
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 225 AATAGAGAGCGCTTACCTTGTACAGCGCTTACGTTATGGAATATGCGCGGTGGCA 284
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 685 ArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProProGluThr 704
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 285 TATTTTGGCGACTTCGAGCAAGATACGATAGTTT---ACACAGCGCAT----- 332
QY 705 TrpGlnAspLeuProValAspPheValAsnGlyLysProSerGlnAlaIleuLeuAla 724
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 332 -----*----- 332
QY 725 GlyValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 333 -----ATCAGTGAATAAACCTTA-----CTTCTTATCCAACTGTGGATTAAT 377
QY 745 -----GlnArgTyrSerLeu 749
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 378 CCGTACGATTAAGCTGTCGTCCTTCCCTTCCACTGGCGGATGTCAGAAATTAATCTTT 437
QY 750 GluValIleGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIle 769
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 438 GATTTAAGCAAAAGATTCGCTATCAGAAATCTCTTTATAAAGTCAGACATCAAGC 497
QY 770 SerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlyMetThr 789
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 498 GCGTGATTCGTATTCAGAAAT-----CATGATATGTT 536
QY 790 GlyGlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArgLeu 809
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 537 GCTCGCTGCAATCGGATGATTTACTATCAAAAGATATGAAAAATTCGCCCTACAGT 596
QY 810 ArgPhePheAlaGlyLysAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerPro 829
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 597 CGTTCTTTCGTCGTCGTCATGCGACGTGCGGTGTCGCGGTATTAATAAATTCGCCCT 656
QY 830 IleSerAspLyGlyTyrLeuThrGlyGlyGlnValIleuAlaValGlyThrAlaGlyIyr 849
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 657 AAAAATAAAATGGAAATTCGGGTGCTGCTACGTTTCTTACAGTCTTTAGAAATAT 716
QY 850 AsnTyrGluPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr 869
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 717 CAATATCAAGTTATCCGAAATGTCGGTGCACATTTTGCAGATAGTGATTAAGCCGCT 776
QY 870 AspLyGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTyrAlaSer 889
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 777 AATAT--TACACAGCAAAAGAGCTCGTTATGCGCAGCGCTTGTGCTTGGCATCG 835
QY 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlyGluGlyAsnProIle 909
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 836 CCAAGTGGCGCGATTAATTTGATATGCGACACCCATTCGCGATTAAGATACAGCAAA 895
QY 910 LysLeuHisPhePheIleGly 916
   :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 896 AATATTCATTTTACATTCGA 916

```

```

RESULT 13
US-10-148-534-6
; Sequence 6, Application US/10148534
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: STANS INSTITUTT FOR FOLKEHELSE
; TITLE OF INVENTION: 85kDa MEISSERIAL ANTIGEN
; FILE REFERENCE: P0235270
; CURRENT APPLICATION NUMBER: US/10/148,534
; CURRENT FILING DATE: 2002-10-09

```

```

; PRIOR APPLICATION NUMBER: GB-9928197.4
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: GB-0005698.6
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: SeqWing9, version 1.02
; SEQ ID NO: 6
; LENGTH: 2379
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
US-10-148-534-6

Alignment Scores:
Pred. No.: 1,29e-11      Length: 2379
Score: 265.50           Matches: 195
Percent Similarity: 34.28% Conservative: 120
Best Local Similarity: 21.22% Mismatches: 348
Query Match: 5.62%      Indels: 257
DB: 9                  Gaps: 39

US-09-914-168-2 (1-919) x US-10-148-534-6 (1-2379)
OY 116 ThrProLeuSerLeuGlulLeuPheAlaGlnGluSerThrGluMetGlyIleAsnPro 135
   ::::::::::::::::::::
DB 46 TCGCCTTTGGCATTTTCCGACCTTCACCATTCAGACATCCGTCTCGACGCTTGCAGCGT 105
   ::::::::::::::::::::
OY 136 AsnAsp-----TyrIleProGluTyrGlnGlyGlnGluInProAsnSer 149
   ::::::::::::::::::::
DB 106 ACCGACCGCAGCACCTTATTCACACTGCGCCGTCAAAGTGGCGACACCTTACAC--- 162
   ::::::::::::::::::::
OY 150 GluValValAlaProProThrLeuGluProGluLysProGlyLeuIleuLysArgLeuTyr 169
   ::::::::::::::::::::
DB 163 -----GACACACACGCGACGTGCCATCATCAAAACCGTGTAC 198
   ::::::::::::::::::::
OY 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
   ::::::::::::::::::::
DB 199 GCCACCGCTTTCTTTGACGACGTCGAAAGTCGACGCGACGCGACGCTTGTGTGACC 258
   ::::::::::::::::::::
OY 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThr 206
   ::::::::::::::::::::
DB 259 GTTATTCGAACGCCGCCACATCGCTCGCTCACATCACCGCGCCCAAAATGCTGCACAAAC 318
   ::::::::::::::::::::
OY 207 GluProTyrAlaAsnIleLysAlaAlaLeuGluAsp-----IleThrGlnGluSerAla 224
   ::::::::::::::::::::
DB 319 GACGCC-----ATCAAGAAAAACCTCGATCGTTCGGCTGCGCGACGTGCAATAC 369
   ::::::::::::::::::::
OY 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
   ::::::::::::::::::::
DB 370 TTTAATCAGCGCAGACTCAACGACGAGTCGCGCGCTGAAAGCAAAATACCTCGGCGCT 429
   ::::::::::::::::::::
OY 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
   ::::::::::::::::::::
DB 430 GCGAAATCAATATTCCAATTCACGCGCCCAAGTACCAAACTCGCCCGCAACCGCGTGAC 489
   ::::::::::::::::::::
OY 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
   ::::::::::::::::::::
DB 490 ATCGACATCATCATTCGACAGGCGCAATATCGCCCAAAATCACCGCATTCGAATTTGAA--- 546
   ::::::::::::::::::::
OY 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289
   ::::::::::::::::::::
DB 547 GGCACCAACCAAGTCTATTCGACCGCAACCAAGTATCGG-----CAG 585
   ::::::::::::::::::::
OY 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisHis 309
   ::::::::::::::::::::
DB 586 ATGTGCTACCGCAAGCGCGCATTTGGACATGCTGACACGAAAGCAGCGGTTGCACCGC 645
   ::::::::::::::::::::
OY 310 GlyLysTyrGluThrLysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPhe 329
   ::::::::::::::::::::
DB 646 CAGAATATTCGCCAGACATGGAAGAAAGTAAACGACTTCTACACAGAAACAGCGCTACTTC 705
   ::::::::::::::::::::
OY 330 AspGlyArgTyrLeuAspArgSerValAspValIleLeuProAspAsnThrAlaAspVal 349
   ::::::::::::::::::::
DB 706 GATTTTCGTAATCCTCGATACCGACATCCAAACCAACGAAAGCAAAACAGGCGACCATTC 765

```

```

OY 350 SerLeuIleTyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIle 369
   ::::::::::::::::::::
DB 766 AAATATCACCTCCACGAAGCGGACGCTTTCGCTGGGCGAAAGT-----TCGATT 816
   ::::::::::::::::::::
OY 370 AspProLysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeu 389
   ::::::::::::::::::::
DB 817 GAAGCGCAGACCAACGAA-----GTCCCAAGGCCGGA 849
   ::::::::::::::::::::
OY 390 LeuGlnGlnLeuThrValAlaAsnMetGlyGluAlaTyrTrsLeuGlnGluValAlaGla 409
   ::::::::::::::::::::
DB 850 CTGGAAAAACTGCTGACACGAAGCCCGCAAGTGTATCAGAACCGCAGATGACCGCC 909
   ::::::::::::::::::::
OY 410 LeuSerTrsAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPhe 429
   ::::::::::::::::::::
DB 910 GTTTGGGT----- 918
   ::::::::::::::::::::
OY 430 ProGluArgGluGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArg 449
   ::::::::::::::::::::
DB 919 -----GAGATTGCAACCCGATGGCTCGGACGCTACGCATACAGC----- 960
   ::::::::::::::::::::
OY 450 ThrGluProAlaGlnValAspGluSerThrLeuGluProValIleGluThrValGluLeu 469
   ::::::::::::::::::::
DB 961 -----GAAATCAGCGTACAGCGCTGCCGGAACCGCGGAACCCAA 999
   ::::::::::::::::::::
OY 470 ThrAspGlyIleLeuMetAspIleSerPro-----IleGlu 481
   ::::::::::::::::::::
DB 1000 ACCGTGATTTGCTCTCGTCGACATCGAACCGCGCGGAAATCTACGTCACGAATTCAC 1059
   ::::::::::::::::::::
OY 482 PheSerAlaSerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHis 501
   ::::::::::::::::::::
DB 1060 ATCACCAGGCAACAAACAAACCCGCGACGAA-----GTCTGGCGCGCGAATTCGCGCA 1113
   ::::::::::::::::::::
OY 502 LeuTyrAspMetProAspAsp-----ArgValLeuAla 512
   ::::::::::::::::::::
DB 1114 ATGGAATTCGCGCGCTTACGACACCTCCAAACGTCGACGTCGAAAGACCGCTGAGCTT 1173
   ::::::::::::::::::::
OY 513 IleAsnHisAspAspGlyValAlaAsnArgSerIleLeuGlyArgIleSerAspAlaValSer 532
   ::::::::::::::::::::
DB 1174 TTGCGCTACTTCCGACCAACGTCACAG-----TTTGAT 1203
   ::::::::::::::::::::
OY 533 AlaValAlaArgAlaIleLeuProAspGluSerGluAsnGluValIleAspLeuProGlu 552
   ::::::::::::::::::::
DB 1204 GCCGTCGCGCTGCGGTACGCGCGACGCAAA-----GTGCGATTGG----- 1242
   ::::::::::::::::::::
OY 553 ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValPro 572
   ::::::::::::::::::::
DB 1243 AACATGAGCCTGACCGACGCTCCACC----- 1269
   ::::::::::::::::::::
OY 573 LeuTyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuT-----Gly 590
   ::::::::::::::::::::
DB 1270 -----GGCTCGCTCGACCTTGACGCGCGC 1293
   ::::::::::::::::::::
OY 591 TrpGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArg 610
   ::::::::::::::::::::
DB 1294 TGGGTTTCAGGATACCGGC-----TTGCTCATTTCCGCGCGCTATTCAGCAGCAACTG 1347
   ::::::::::::::::::::
OY 611 AspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysLysGlyValLysLeu 630
   ::::::::::::::::::::
DB 1348 TTCGGTACGCGGCAAGTCGCGCGCTGCGCGCTCGGGAAGCAAA-----ACCACGCTC 1401
   ::::::::::::::::::::
OY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
   ::::::::::::::::::::
DB 1402 AACGGCTCGGTGCTTACCGACCGCTACTTACGCGCAGACGGGGTACACCTGCGGTAC 1461
   ::::::::::::::::::::
OY 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670
   ::::::::::::::::::::
DB 1462 -----GATATTACGGAAGAGCTTCGACCCCGCAAGCATTCGACACGCTCAAAACA 1515
   ::::::::::::::::::::
OY 671 HisGluIleSerArgSerIleIleGlnAsnGlyGlyTyrAsnArgThrTyrSerLeuArg 690
   ::::::::::::::::::::
DB 1516 TATTAACCAACCAACCGCC-----GGCGGCGCGCTAAGATGGG-TATCCCGCT 1562

```

```

Oy 691 TTTATGLeu-AspLysLeuLysThrGlnAlaProGlu-----ThrT 705
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1563 TACCGAATACGACGGCTCAATTTCGGCGTGGCGGGAACCTGACCTCAACACTA 1622
Oy 705 pGlnAspLeuPro-----ValAspPheValAsn-----GlyLys---ProSerG 719
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1623 CAACAAAGCAACCAACCGATATCCGACTTTATCAGAAATACGCAAAACGACGGCG 1682
Oy 719 nGluAlaLeuAlaGlyAlaValAlaHisLysThrVal-----Al 733
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1683 AGACGGCAGCTTCAAAAGCGCTGTGTAACAAGACCGCTGCGGTGGGGCGCAACAAGAC 1742
Oy 733 aaAspAlaLeuValAsnProMetLarGlyTyrGlnArgLysSerLeuGluValGlyse 753
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1743 CGACACGGCTCTATGGCGCGACGGCGGCTAC----- 1773
Oy 753 rSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyLysSerGlyValTyr 773
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1774 -----CTGACCGGCTTAATGCGCAAAATGCGCTGCGCGCGCAAACTGCATA 1823
Oy 773 rSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGly----- 790
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1824 C-----TACTCGCGCACCCACACCAACCAACCTGTTCTTCC 1859
Oy 791 -----GlyIleGlnAlaGlyTyrIle 797
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1860 CTTAAGCAAAACCTTCACGCTGATCTCGGCGGCAAGTCGGCATTCGGCGCGCTAC-- 1917
Oy 797 eTTPSerAspAsnPheAsnHisValProTyrAlaGlyLeuArgPhePheAlaGlyAspG 817
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1918 -----GCCAGAACCAAGAAATCCCTTTGAAACCTTACGCGCGCGCTGGC 1970
Oy 817 nSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLysGlyTyrLeuTh 837
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1971 TTCCGTCGCGGCTACCAAAACCGGACGCTCGGCGC-----AAAGTATGACGA 2021
Oy 837 rGlyGlyGlnValLeuAlaValGly-----ThrAlaGluTyrAsnTyrGluPheMe 854
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 2022 ATACGGCGAAATATCAGTACGCGCGCAACAAAAAGCCACCTCTCGCGCGAGTGGCT 2081
Oy 854 t-----LysAspLeuArgLeuAlaValPheGlyAspIle 865
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 2082 CTTCGCCATGCCGCTCGCAACAGCAGCAGCAGCAGCAGCTGCGCTGTTGGCGGACCC 2141
Oy 865 eGly----- 866
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 2142 AGCGAGCGTGTGGGAGCGCAACCTATACGCGCGCGCAAAACGTAACAACAATCGGT 2201
Oy 867 -----AsnAlaTyrAspLysGlyPheThrAsnAspThrLysIleGlyAlaGly 883
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 2202 TTACTCGGAAACCGCGCATTAATCCACCTTTACCAACCAATTGGCTATTCGCGCGCGG 2261
Oy 883 LglValArgTyrAlaSerProValGlyGlnValArgValAspValAlaThrGlyVal 903
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 2262 CGCGGTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2321
Oy 903 sGlu-----GluGlyAsnProIleLysLeuHisPhePheIleGlyThrProPhe 919
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 2322 GAAAAACCGGAGACGAAATCCAAACGCTTCCATTCACGCTGCGGACGAGCTTC 2376

```

```

; PRIOR FILING DATE: 1999-11-02
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 2391
; TYPE: DNA
; ORGANISM: Neisseria meningitidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2391)
; OTHER INFORMATION: OMP85 cds
US-10-320-800-5

Alignment Scores:
Pred. No.: 4,42e-11 Length: 2391
Score: 259.00 Matches: 191
Percent Similarity: 33.48% Conservative: 121
Best Local Similarity: 20.49% Mismatches: 342
Query Match: 5.48% Indels: 278
DB: Gaps: 40

US-09-914-168-2 (1-919) x US-10-320-800-5 (1-2391)
Oy 116 ThrProLeuSerLeuGluGlnPheAlaGlnGluSerThrGlnMetGlyIleAsnPro 135
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 46 TCGCTTGGCTGCTGCGGCTTCCATCCAAAGACATCCGCGCTGCAAGGCTTGACGCT 105
Oy 136 AsnAsp-----TyrIleProGluTyrGlnGlyGlnGluProAsnSer 149
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 106 ACGGAGCGGAGTACCGTATTCACACTACCTGCGCTCAAAAGTCGGCGACCTACAAAC-- 162
Oy 150 GluValValValProProThrLeuGluProGluLysProGluLeuIleLysArgLeuTyr 169
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 163 -----GACACACGCGCAGTCCATCAACAACCAACCTGTAC 198
Oy 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 199 GCCACCGGCTTCTTTGACAGCGTACGCGTGGAAACCTGCGGACGCGCTGCTGAC 258
Oy 187 PheTyrGlnSerSerGlnSerGlyLysThrSerAlaIleGlySerSerHisGlnLysThr 206
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 259 GTTATCGAACGCGCCACCATGCGCTGCTGCAACATCAACGCGCAAAATGCTGCAAAAC 318
Oy 207 GluProTyrAlaAsnIleLysAlaIleGluLysPhe-----LleThrGlnLysSerAla 224
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 319 GACGCG-----ATTAAGAAACCTCGAATCGTTCGCGCTGGCGCATGCGCAATAC 369
Oy 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 370 TTATATCAGCGGACACTCAATCAGGAGCTGCGCGCGTGAAGAAATACCTGCGCGCG 429
Oy 237 -----ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAsp 249
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 430 GCGAACTCAATATCAAAATACCGCCCAAGTAAACCAACTGCGCGCGCAACCGGCTGAC 489
Oy 250 IleAspLeuSerIleIleArgAsnSerIleGlyLysValAspValIleIleHisAspLeu 269
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 490 ATCGACATCAGATTCAGCGAGGCGCAATCCGCGCAAAATCAGCGCATGGAATTGAA-- 546
Oy 270 GlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 547 GCGAACCAAGCTATTCGACCGCGCAACTGATGCGG----- 582
Oy 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 583 -----CAAAATGCCCTGACCGGAGCGGATTTGGACATGG 618
Oy 307 -----PheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGlu 320
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 619 CTGACAGCAAGCAACAATTCAGGAGCAAAATTTGCCAAGATATGAAAAAGTAAACC 678
Oy 321 AsnLaserAlaGluHisGlyTyrPheAspGlyArgTyrPheAspArgSerValAspVal 340
      |||||:||||:||||:||||:||||:||||:||||:||||:||||:

```


Db 679 GACTTCTACCAAAATACGGCTACTTCGATTTCCTCGATCCGACATCCAAACC 738
 QY 341 ILeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrIleTyrArg 360
 Db 739 AACGAAGACAAAACAGACGACCATCAAAATACCGTCCACGAGGCGGCGCTTCCGT 798
 QY 361 pheAspGluValValPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
 Db 799 TGGGGCAAAAGTC-----TCCATCGAAGGCGACACCAAGAA----- 834
 QY 381 AspLysLeuProValLysArgGluLeuLeuGlnLeuLeuThrValAsnMetGlyGlu 400
 Db 835 -----GTCCCAAAAGCCGAACCTGGAAAACTGCTGACCATGAAAGCCCGGCAAA 882
 QY 401 AlaTyrAsnLeuGlnAlaValArgAlaIleSerAsnAspLeuLeuAlaThrArgTyrPhe 420
 Db 883 TGGTACGAACGCCAGACGATGACCGCGTTTGGT----- 918
 QY 421 AsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGlnAsnAspGlnVal 440
 Db 919 -----GAGATTCGAAGCCCGCATGGGC 939
 QY 441 SerPheGlnGlnSerSerSerArgThrGluProAlaGlnValAspGlnSerThrLeu 460
 Db 940 TCGGCGAGCTACGCGCATACGC-----GAAATCAGCGCTA 972
 QY 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetLysIleSerPro--- 479
 Db 973 CAGCCGCTGCGAAGCCTGAACCCAAACCGTGATTTCTGCTGCACATCGAAGCGGCG 1032
 QY 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspLysLeu 492
 Db 1033 CGGAAATGTACGTCAACGAAATACATCACCGCGCAACAAACCCGCGAGAA--- 1089
 QY 493 AsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAsp--- 508
 Db 1090 ---GTCTCCGCGCTGAATTAATTCGCAAAATGGAATCCGCGCATTAAGACCTTCAAGCTG 1146
 QY 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle 523
 Db 1147 CAACGTTCCAAAGAGCGCGTTCGAGCTTTGGGTACTCTGCAAAATGCTCCAG--- 1197
 QY 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGlnSer 543
 Db 1198 -----TTTGATGCTGTCCCGCTTCCGCGCGCGCGCAAAA--- 1233
 QY 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
 Db 1234 -----GTGATTTTG-----AACATGAGTCTGACCGAAGCTTCCACC----- 1269
 QY 564 AspValTyrGlnSerLysValProLeuTyrValPheValAlaSerAspLysProArg 583
 Db 1269 ----- 1269
 QY 584 AspGlyGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr 601
 Db 1270 ---GGTTCCTGGATTGTAGCGCGGGGTGGGTTCAGATATCCGCG-----TTGGTCATG 1320
 QY 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaLeuAlaGluLeuArgLeu 621
 Db 1321 TCCGCGAGCGCTTCCCAAGCAACCTGTCGGGTACGCGCAAGTCCGCGCGCGCGGCC 1380
 QY 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
 Db 1381 TCCAGAGAGCAAA-----ACCAGCTTAACGGGTCCGCTTACTGACCCCTACTTC 1434
 QY 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGly 661
 Db 1435 ACGGCGAGCGGGGTGACGCGGCTAC-----GATGTTTACGAAAAGCCTTCGACCCG 1488
 QY 662 PheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGly 681
 Db 1489 CGCAAAAGCATCGACGATCAAAACAATTAATAACACACCGCGCA-----GGC 1536

QY 682 GlyTyrAsnArg-----ThrTyrSerLeuArgTyrArgLeuAspLysLeu 696
 Db 1537 GCAGGATCCGCGATGAGCGTGGCTGTACCGAATACGACCGCGGTGATTTGGTTGGTG 1596
 QY 697 LysThrGlnAlaProProGluThrTyrGlnAspLeuPro-----ValAspPheVal 713
 Db 1597 GCAGAACCTGACCGCTCAACACCTTACCAACAAAGCGCCCAAAACATATGCGCACTTATC 1656
 QY 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730
 Db 1657 AMGAATACGGCAAAACCGACGCGACAGCGAGCTTCAAAAGCGCGGTGTCAAAAGT 1716
 QY 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
 Db 1717 ACCGTGGCGTGGGGCGGCAACAAACCGACGCGCTTATGGCGGCGCGCTAC--- 1773
 QY 745 GluArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
 Db 1774 -----CTGACGGCGGTGAACCGCGCAATC 1797
 QY 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
 Db 1798 GCCCTGCTGGCGAGCAACTGCAATAC-----TACTCGCC 1833
 QY 785 AlaHisGlnMetThrGly----- 790
 Db 1834 ACCCACAACCAACCTGTTCTTCCCTGAGCAAAACCTTCAAGCTGATGCTGGCGGC 1893
 QY 791 -----GlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArg 808
 Db 1894 GAAGTCGGCATTCGCGGCGCTAC-----GGCAGAACCAAAAGAACCTTCTTT 1944
 QY 809 LeuArgPhePheAlaGlyGlyAspArgGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
 Db 1945 GAAACTCTACGCGCGCGCGCTGGTGGTGGCGGATGACAAAGCGGACGCTCGGT 2004
 QY 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGly----- 845
 Db 2005 CCG-----AAAGTCTATGACGAATACGGGCAAAATTCAGCTACGGCGGCAACAA 2055
 QY 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
 Db 2056 AAAGCCACAGCTTCGCGCGCGCTCTTCCGATGCCCGCGCGCAAGACGCGCGCAC 2115
 QY 857 LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
 Db 2116 GTCCGCTGAGCTGTTTGGCGGACGAGCGAGCTGTGGGACGCAAAACCTACGACGAC 2175
 QY 870 ----- 870
 Db 2176 AACAGCAGTCCGCGACCGCGCGGAGGTTCAAAACATTTACGCGCGCGCAATACCAT 2235
 QY 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTyrAlaSer 889
 Db 2236 AAATCCACCTTTATCCAAAGAAATGCGCTATTCGCGCGCGCGCTGCTGCTCG 2295
 QY 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
 Db 2296 CCTTTAGGCGCGGAAATTCACCTACCGCTGAAAGAAAAACCGGAAGAGAA 2355
 QY 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
 Db 2356 ATCCAACGCTTCCAAATTCCAACTCGCGACGACGCTC 2391

RESULT 15
 US-10-148-534-2
 ; Sequence 2, Application us/10148534
 ; GENERAL INFORMATION:
 ; APPLICANT: CHIRON SPA
 ; APPLICANT: STRATENS INSTITUIT FOR FOLKEHELSE
 ; TITLE OF INVENTION: 85KDA NEISSERIAL ANTIGEN
 ; FILE REFERENCE: P0235270


```

Db 1489 CGCAAGCATCGACGACATCAAAATATMAAACGACGACGCA-----GGC 1536
Oy 682 GLYTPASArG-----ThrTyrSerLeuArgTyrTrgLeuAspLysLeu 696
Db 1537 GCAAGCATCCGCGATGAGCTGCTGTACGAATACGACCGCGTGAATTTGGTTGGTG 1596
Oy 697 LysThrGlnAlaProProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713
Db 1597 GCAGAACACCTGACCTCAACACCTTACAAAGCGCCCAAAACATGATGCCGACTTATC 1656
Oy 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyAlaValHisLys 730
Db 1657 AAGAAATACGCGCAAAACCGACGCGACGCGACGCTTCAAAAGCTGTGTACAAAGT 1716
Oy 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgGlyTyrArg 744
Db 1717 ACCGTGGCTGGGGGGCGCAACAAACCGACGCGCTTATGCGCGACGCGCGCTAC--- 1773
Oy 745 GlnArgTyrSerLeuGlnValGlySerSerGlyLeuValSerAspAlaAsnMetAlaLe 764
Db 1774 -----CTGACGCGCGGTGAACGCGCAATC 1797
Oy 765 AlaArgAlaGlyLysSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
Db 1798 GCCCTGCTGCGCAAACTGCAATAC-----TACTCCGCC 1833
Oy 785 AlaHisGlnMetThrGly----- 790
Db 1834 ACCCAACACCAACCTGCTTCTTCCCTGAGCAAAACCTTCACGCTGATGCTCGGCGC 1893
Oy 791 -----GlyLleGlnAlaGlyTyrLleTrpSerAspAsnPheAsnHisValProTyrArg 808
Db 1894 GAAGTGGCATTCGCGCGCGCTAC-----GGCAGAACCAAGAAATCCCTTCTTT 1944
Oy 809 LeuArgPhePheAlaGlyLysAspGlnSerLleArgGlyTyrLleHisAspSerLeuSer 828
Db 1945 GAAACTTCTACGCGCGCGCTGCTGCTGCGGATACGAAAGCGGACGCTCGGT 2004
Oy 829 ProLleSerAspLysGlyTyrLleuThrGlyGlyGlnValLeuValAlaGly----- 845
Db 2005 CCG-----AACTCTATGACGAATACGCGGAAATAATCATCGCGCGCAACAA 2055
Oy 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
Db 2056 AAGCCACAGCTCTCCGCGCGCTCTTCCGATGCCGCGCGCAAGACGCGCGAC 2115
Oy 857 LeuArgLeuAlaValPheGlyAspLleGlyAsnAlaTyrAsp----- 870
Db 2116 GTCCGCTGAGCTGTGTTGCCGACGACGCGTGTGGGCGCAAAACCTACGACGAC 2175
Oy 870 ----- 870
Db 2176 AACAGCAGTTCGCGACGCGCGCGAGGTTCAAAACATTTACGCGCGCATACCCAT 2235
Oy 871 ---LysGlyPheThrAsnAspThrLysLleGlyAlaGlyValAlaArgTyrPalaSer 889
Db 2236 AATTCACCTTACCAACGATTCGCTATTCGCGCGCGCGGTACTGCTCTCG 2295
Oy 890 ProValGlyGlnValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
Db 2296 CTTTAGGCCCGGAAATTCACCTACCTACCGCTGAAGAAAAAACCGGAAGACGAA 2355
Oy 908 ProLleLysLeuHisPhePheLleGlyThrProPhe 919
Db 2356 ATCCACAGCTTCATTCCAATTCGCGACGACGCTC 2391

```

```

; FILE REFERENCE: P02352720
; CURRENT APPLICATION NUMBER: US/10/148,534
; CURRENT FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: GB-9928197.4
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: GB-0005698.6
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 106
; SOFTWARE: SeqMan99, version 1.02
; SEQ ID NO 10
; LENGTH: 2394
; TYPE: DNA
; ORGANISM: Neisseria gonorrhoeae
US-10-148-534-10

Alignment Scores:
Score: 4.43e-11 Length: 2394
Percent Similarity: 259.00 Matches: 191
Best Local Similarity: 33.48% Conservative: 121
Query Match: 20.49% Mismatches: 342
9 Indels: 278
Gaps: 40

US-09-914-168-2 (1-919) x US-10-148-534-10 (1-2394)
Oy 116 ThrProLeuSerLeuGlnGluPheAlaGlnGluSerThrGlnMetGlyLleAsnPro 135
Db 46 TCGCTTTGGCACTTCCGCGACTTCACCATTCAGACATCCGCGCTCGAAGCTTGACGCT 105
Oy 136 AsnAsp-----TyrLleProGluTyrGlnGlyGlnGlnProAsnSer 149
Db 106 ACCGACCGCAGTACGCTATTCACACTCCGCGCAAAAGTCCGCGCACCTCAAC--- 162
Oy 150 GlnValValAlaProProThrLeuGlnProGlnLysProGlyLysLeuLysArgLeuTyr 169
Db 163 -----GACACACAGCGCAGTCCATCATCAAAACCTGTAC 198
Oy 170 AlaArgLeuPheAsnAspGlyValAsnLysValPro-----ArgLeuLysAlaLys 186
Db 199 GCCACGGTTCTTTGACGACGATACCGGTGCAAACTGCGGACGCGGACGCTCGTGAC 258
Oy 187 PheTyrGlnSerSerGlnSerGlyGluThrSerAlaLleGlySerSerHisGlnLysThr 206
Db 259 GTTATGACAGCCGCCACCATCGCTGCTCAACATCACCGCGCAAAATGCTGCAAAC 318
Oy 207 GluProTyrLalaSnLleLysAlaLalaLeuGluAsp-----LleThrGlnGluSerAla 224
Db 319 GACGCC-----ATTAGAAAAACCTCGAATCGTTCGGCGTGGCGCAGTCGCAATAC 369
Oy 225 MetAsp-----LeuAsnGlySerLleProArgLeuArgGln----- 236
Db 370 TTTAATCAGCGCACATTCATGACGATCGCGCGCTGAAAGAAATATCTCGGGCGC 429
Oy 237 -----ThrAlaLeuValAlaAlaArgAlaValAlaGlyTyrTyrAsp 249
Db 430 GCCAAATCAATATCCAAATACGCGCCAAAGTACCAAACTCCGCGCAACCCGCTGAC 489
Oy 250 IleAspLeuSerLleLleArgAsnSerLleGlyGlnValAspValLleLleHisAspLeu 269
Db 490 ATCGACATTCAGATTCGACGAGGCGCAATCCGCCAAATATCACCGACATCGAATTTGAA--- 546
Oy 270 GlyLnuProValTyrLleAspTyrArgAlaValAlaGlnValArgLysGlnGlyAlaAspAsp 289
Db 547 GGCACACAGTCTATTCGACCGCAAACTGATGGG----- 582
Oy 290 LysAlaPheThrThrValAlaAspGluValProLeuLeuLleGlyAspVal----- 306
Db 583 -----CAGATGCTCCTGACCGCAAGCGCGCATTTGACATGG 618
Oy 307 -----PheHisHisGlyLysTyrGluThrLysLysAsnLeuLleGlu 320
Db 619 CTGACACGAAGCAACCAATTCAGACGAGAAATTTGCCCAAGACATGAAAAAGTAACC 678

```

QY 321 AsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgThrLeuAspArgSerValAspVal 340
 Db 679 GACTTCTACCAAGACAGACGGCTACTTTCGATTCCTCGATCCGACATCCAAACC 738
 QY 341 IleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArg 360
 Db 739 ACGAAGACAAACCAAGACGACCATCAAAATCACCTCCACAGACGGCGGCTTCCTCGT 798
 QY 361 PheAspGluValValIlePheThrThrIleAspProLysThrAsnGlnLeuThrThrAspPro 380
 Db 799 TGGGGCAAAAGTC-----TCCATCAAGGCGACACCAAGAA----- 834
 QY 381 AspIleLeuProValIleTyrArgGluLeuLeuGlnLeuLeuThrValIleAsnMetGlyGlu 400
 Db 835 -----GTCCCAAGACCGCAACTGGTGAACCATGAAGACCGCGGCAAA 882
 QY 401 AlaTyrAsnLeuGlnAlaValIleArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
 Db 883 TGGTACCAAGACCGCAGATGACCGCGCTTTGGGT----- 918
 QY 421 AsnMetValAsnThrGluIleValIlePheProGluArgGlnIleGlnAsnAspGlnVal 440
 Db 919 -----GAGATTCAGAAACCGCATGGGC 939
 QY 441 SerPheGlnGlnSerSerSerArgThrGluProAlaGlnValAspGluSerThrLeu 460
 Db 940 TCGGACCGCTACGCATACAGC-----GAATACAGCGTA 972
 QY 461 GluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro--- 479
 Db 973 CACGCGCTGCCCAACGCCGCAAAACCGTCGATTTCCGTCGACATCGAAGACGGGC 1032
 QY 480 -----IleGluPheSerAlaSerAsnLeuIleGlnAspIleLeu 492
 Db 1033 CGGAATATCTACGTCACAAAGAAATCCACATCCACGGCAACAAACCGCGGAGAA--- 1089
 QY 493 AsnLeuValAlaAlaIleValArgHisLeuTyrAspMetProAspAsp--- 508
 Db 1090 ---GTCTGCGCCCGCGCAATGGAATCGCGGCTTACGACACTCCAAAGCTG 1146
 QY 509 -----ArgValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIle 523
 Db 1147 CAACGCTCCAAAGACGGCTGAGCTTTGGGTACTTCGCAACGTCACG----- 1197
 QY 524 LeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAspGluSer 543
 Db 1198 -----TTTGATGCCGTCGCCGCTTCCGCGCACACCGGACAAA--- 1233
 QY 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
 Db 1234 -----GTGCAATTTG---AACATGAGCCTGACCCGAAACGTTCCACC----- 1269
 QY 564 AspValTyrGlnSerLysLysValIleProLeuTyrValIlePheValAlaSerAspLysProArg 583
 Db 1269 ----- 1269
 QY 584 AspGlyGlnIleGlyLeu-----GlyTyrGlySerAspThrGlyThrArgLeuValThr 601
 Db 1270 ---GGCTCGCTCGACTTGACGGCGGCTGGTACAGATACCGGC-----CTGGTCAAG 1320
 QY 602 LysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGlnLeuArgLeu 621
 Db 1321 TCCGACGCGCTTCCCAAGACAACTGTCGTTACGGGCAAGTCGCGCCCTGCGCGCC 1380
 QY 622 SerGluAspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsn 641
 Db 1381 TCACCAAGCAAA-----ACCAGGCTCAAGCGCTGCTGTGTACCGACCGCTACTTC 1434
 QY 642 AspGlnLeuArgAlaThrLeuGlyTyrGlnGlnGluValIlePheGlyHisSerThrAsnGly 661
 Db 1435 ACGGACGACGGGCTGACGCTGGGTAC-----GATGTTTACGGAAGACCTTCGACCGG 1488
 QY 662 PheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSerIleIleGlnAsnGly 681

Db 1489 CGCAAGCATCGACACGACATCAAAATATAAACCACCGCA-----GGC 1536
 QY 682 GlyTyrAsnArg-----ThrTyrSerLeuArgTyrIleLeuAspLysLeu 696
 Db 1537 GCAGGACATCGCATGAGCGCTGTTCCGATACGACCGCGCTGAATTTCCGTTGGGTG 1596
 QY 697 LysThrGlnAlaProProGluThrTyrGlnAspLeuPro-----ValAspPheVal 713
 Db 1597 GCAGAACACCTGACCGCATACACCTTACAAACAGCGCCAAACGATATGCGGACTTATC 1656
 QY 714 Asn-----GlyLys---ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys 730
 Db 1657 AAGAAATACCGCAAAACCGACGCGACAGCGGAGCTTCAAGCGTGGCTGTACAAAGT 1716
 QY 731 ThrVal-----AlaAspAsnLeuValAsnProMetArgIleTyrArg 744
 Db 1717 ACCGTGCGCTGGGGCGCCCAACAAACCGACGCGCTTATGGCGGACCGCGGCTAC--- 1773
 QY 745 GlnArgTyrSerLeuGluValIleGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
 Db 1774 -----CTGACGGGCGCTGAACGCGCAAAATC 1797
 QY 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
 Db 1798 GCCCTGCCGCGACGCAAACTGCAATAC-----TACTCGCC 1833
 QY 785 AlaHisGlnMetThrGly----- 790
 Db 1834 ACCCAACCAACCAACCTGTTCTCCCTTAAGCAAAACCTTACACCGTATCGTCGGCGGC 1893
 QY 791 -----GlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArg 808
 Db 1894 GAATCGGCATTTGGCGGCGGTAC-----GCGAACAACAAAGAAATCCCTTTT 1944
 QY 809 LeuArgPhePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
 Db 1945 GAATCTTCTACGGCGGCGGCTGGTGGTGGCGGATACGAACCGGACGCGCTCGT 2004
 QY 829 ProIleSerAspLysGlyTyrLeuThrGlyGlnValLeuAlaValGly----- 845
 Db 2005 CCG-----AAAGTGTATGACGAATACGCGCAAAATACACTACGGCGGGAACAA 2055
 QY 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
 Db 2056 AAAGCCAACTCTCGCGGACGCTCTTCCGATGCCCGGCGGAAAGACGGCGGACCC 2115
 QY 857 LeuArgLeuAlaValIlePheGlyAspIleGlyAsnAlaTyrAsp----- 870
 Db 2116 GTCCGCGTGAACCTGTTGCCGACGACGACGCGTGTGGGACGCAAAACCTACGACGAC 2175
 QY 870 ----- 870
 Db 2176 AACAGCACTTCCGCGACCGCGGCGGCAAGGTTCAAAACATTTCAGGCGCGGCAATACCAT 2235
 QY 871 ---LysGlyPheThrAspAspThrLysIleGlyAlaGlyValIleValArgTyrAlaSer 889
 Db 2236 AAATCCACCTTACCAAGAAATTCGCGATTCGCGCGGCGGCGGTACTGCTGCTCG 2295
 QY 890 ProValIleGlyGlnValArgValAspValAlaThrGlyValLysGly-----GluGlyAsn 907
 Db 2296 CTTTAGGCGCGGATGAATTCAGTACGCTCAACCGCTGGAAGAAACCGAAGACGAA 2355
 QY 908 ProIleLeuHisPhePheIleGlyThrProPhe 919
 Db 2356 ATCCAAACGCTTCCAAATTCCAACTGCGGACGACGTTTC 2391
 RESULT 17
 US-09-689-065A-1
 ; Sequence 1, Application US/09689065A
 ; GENERAL INFORMATION:
 ; APPLICANT: Pfizer Products, Inc.
 ; TITLE OF INVENTION: LAMSONIA INTRACELLULARIS PROTEINS AND RELATED METHODS AND

```

: TITLE OF INVENTION: MATERIALS
: FILE REFERENCE: 3153, 00187/PC10589A
: CURRENT APPLICATION NUMBER: US/09/689,065A
: PRIOR FILING DATE: 2000-10-12
: PRIOR APPLICATION NUMBER: US Prov. 60/160,922
: PRIOR FILING DATE: 1999-10-22
: PRIOR APPLICATION NUMBER: US Prov. 60/163,858
: NUMBER OF SEQ ID NOS: 112
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO: 1
: LENGTH: 6617
: TYPE: DNA
: ORGANISM: Lawsonia intracellularis
US-09-689-065A-1

Alignment Scores:
Pred. No.: 7,85e-09 Length: 6617
Score: 238.50 Matches: 198
Percent Similarity: 34.72% Conservative: 137
Best Local Similarity: 20.52% Mismatches: 391
Query Match: 5.05% Indels: 239
DB: 7 Gaps: 38

US-09-914-168-2 (1-919) x US-09-689-065A-1 (1-6617)
Oy 50 AsnGlnAlaIyAlaGlyAsnProProValLeuLeuThrProGluGlnIleGlnAlaArg 69
Db 3764 AATGCTGCTTCAAAAGACGATCTTCTATTTGGTTCTCCCAATTCAAATTAATGGCTCA 3823
Oy 70 LeuAsnAlaIaGlyLeuAsnAlaIyS---ProGlnSerGlnAlaLeuAspValAlaAsn 88
Db 3824 TCAATGATGAAAGATTTCACAAACAGACACCAATGCTCTTGCAACGATTAAGAAT 3883
Oy 89 PheAspAspGlnSerProIleSerArgIleGlyGluGlnSerProProLeuGlyLeuAsp 108
Db 3884 -----AAGGCA 3889
Oy 109 MetSerValIleGluGluThrThrProLeuSerLeuGluLeuPheAlaGlnGlnSer 128
Db 3890 TTTCGTGTCATCCCTAATTAATCTGCATTAACTTT-----CTATATAACAAAATATAC 3943
Oy 129 ThrGluMetGlyIleAsnProAsnAspTyrIleProGluTyrGlnGlyGluGlnProAsn 148
Db 3944 TCCCAACTTAATATTTCTACTGCAAAAAGGTA-----GCTCAACAACCTCCAT 3991
Oy 149 SerGluValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeu 168
Db 3992 GCTGACTATGTAGTA----- 4006
Oy 169 TyrAlaIaArgLeuPheAsnAspGlyValAlaLysValProArgLeuLysAlaLysPheTyr 188
Db 4007 TACGGCAGTTTCAATCAACACAGGTGAAT-----TTTACT 4042
Oy 189 GlnSerSerGlnSerGlyLysThrSerAlaIleGlySerSer-----HisGlnLys 205
Db 4043 ATTGATAGTAGCCTTATTCATGTACAGGTGTACGATCTGCACGCCCATTTATACATAGAA 4102
Oy 206 ThrGluProTyrAlaAsnIleLysAlaIaLeuGluAspIleThrGlnGlnSerAlaMet 225
Db 4103 AAACCAAAATTTAATGAGCTAATATTTGCTGTACAGAACTGCTGCAAGCTAATAGT--- 4159
Oy 226 AspLeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaIaIaArgAlaVal 245
Db 4160 -----AATGCCCTTATTAAGAAA---AACACTATTGCTGATGTACGATTTCATGGGCTT 4210
Oy 246 GlyTyrTyrAspIleAspSerIleIleArgAsnSerIleGlyGluValAspValIle 265
Db 4211 AAAGTTCTTGATCTGATGTATCTTACACAGCTCACATTAATTAAGGAGATCTACTACT 4270
Oy 266 IleHis-----AspLeuGlyLysProValTyr 274
Db 4271 GATCATGCCCAAAATTAATGTCAGAAATCAAAAATAATGCGAATTGGA-----TATTTT 4324

```

```

Oy 275 IleAspTyrArgAla---ValGluValArgIleGlyGluValAlaAspAspLysAlaPheThr 293
Db 4325 AGTGATGCTCTGCGCAAGTATTAAGAAACGGGGAAGA-----CGATTACTTGTA 4375
Oy 294 ThrValAlaAspGluValProLeuLeuIleGlyAspValPheHisIleGlyLysTyrGlu 313
Db 4376 TTACTGTACAGAAAGGCTAAATTTACAGATGTTGTTGTCACAGGCTCAAAAGCTGTA 4435
Oy 314 ThrLysLysAsnLeuIleGluAsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgTyr 333
Db 4436 ACTATGATAAACATCTTGCTGCAATGAGTCTAAAGGA----- 4477
Oy 334 LeuAspArgSerVal-----AspValIleLeuProAspAsnThrAlaAspValSerLeu 351
Db 4478 -----TCAGTTATTAGTAGTACCTATTGTCCTCAAGATATTCAAAAATTTACCGAC 4528
Oy 352 IleTyrAspThrGlyThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspPro 371
Db 4529 CTCTATAGAAAAGAGGCTACTATCTCGCTGAA-----GTTAATTATGAATTAAGAG 4582
Oy 372 LysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGlu 391
Db 4583 AAAGAAAATATCTTCTGTGCA-----ACA 4606
Oy 392 GlnLeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSer 411
Db 4607 CTATTGTTACAGTAATATGAAGGAAACCTTTATATTAAGATGTCGGAATTGACAGA 4666
Oy 412 AsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGlu 431
Db 4667 CTGGAACATATAAGCTTAA-----ACTTTAAAAAAGAGTTACATTAAACAGA 4717
Oy 432 Arg-----GluGlnIleGlnAsn 437
Db 4718 CGTAATTTTATTCATGTTTACTGACAGGCTGATTAACGTGAAGAAATCTTGACACT 4777
Oy 438 AspGlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnIleAspGlu 457
Db 4778 GACTCTATAGCATC-----TCTGCTATGCCATTAATCATGCGTATAGTATAT 4828
Oy 458 SerThrLeuGluPro-----ValIleGluThrValGluLeu 469
Db 4829 CAAGTTGCTTCCACCGAATACATTCATGAAAGAAAGATTTGTTATTCATAGAGTA 4888
Oy 470 ThrAspGlyIleLeuMetLysPheSerProIleGluPheSerAlaSerAsnLeuIleGln 489
Db 4889 AAGAAGGTAAAGCCCTATTAATAGCAAAAATAGACTTTAAAGAGACTTATTTAGACA 4948
Oy 490 AspLysLeuAsnLeuValAlaIaLysAlaArgHisLeuTyrAspMetProAspAspArg 509
Db 4949 AATGAACACATCTCTTAATTAACAAAATTTGATGATCAATTAACATATAGCAGATTTT 5008
Oy 510 ValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAsp 529
Db 5009 TCTCTTCTGTTATGACAGATGAT-----GTAAAAAGCATTAACAGATTTTATTCAGAT 5062
Oy 530 AlaValSerAlaValAlaArgAlaIleLeuProAspGlnSerGluAsnGlu----- 546
Db 5063 TATGTTATGATTTGCTGAAGATAGATCTT---GAACAACCAAAAATGAAGAAGATGCA 5119
Oy 547 -----ValIleAspLeuProGluArgThrAlaLeu----- 556
Db 5120 ACAATGATGATCTTCTTCTTATTTGATTAACAAAAGTCTTCTTCTGTAAGATAATT 5179
Oy 557 -----AlaAsnArgLysThrProAlaAspValTyr----- 566
Db 5180 GTTGAAGAAATATCTGCTACTAGAGATATTTATTCCTCGGAATTAACGCTTGCTGAT 5239
Oy 566 ----- 566
Db 5240 GGAGATCTTTTAAATGTCACACATCTCGACGCTTAATGAATGCCTTAAACGCGCTTGCC 5299

```

```

OY 567 -----GlnSerLysValProLeu 573
Db 5300 TATTTAACCAGTAGATACAGATACAGCTGCTACAGGGAAGATGATGATGATGATCA 5359
OY 574 TyrValAlaHeValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrGlySer 593
Db 5360 CTGTAAAGCTT-----CAAGAAGCTCGAAGACAGTGCATACAGAGTGGTGGTAC 5413
OY 594 AspThrGlyThrArgLeuValThrLys-----PheGlnHisAsnLeuIleAsnArg 610
Db 5414 TCACACACTTCTAAATTGGTGGTTTCAGGAAGTATCTCAGAAAGAACTTATGGGGA 5473
OY 611 AspGlyTyrGlnAlaGlyAlaGlyLeuArgLeuSerGlyAspLysGlyValLysLeu 630
Db 5474 ---GGTATATTTAAAGTATGAGGTTTATTTCTAGTAAGTCACTCTCTTGATCTT 5530
OY 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
Db 5531 TCTTTTACCACATCCCTGCTTATGATACAGAC----- 5563
OY 651 GlnGlnGlnValPheGlyHisSerThrAsnGlyPheAspLeu----- 664
Db 5564 -----TTTGCGCTTATGATATACATTTATACGCTACGAGATGATGGATGAC 5611
OY 665 -----SerThrArgThrLeuGlnHisGlyIleSerArgSerIle 677
Db 5612 TTCGCTAAAGAACTTATGAGATACACATACGCTATTTCCACCTATA----- 5659
OY 678 IleGlnAsnGlyTyrPheAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLys 697
Db 5660 -----GGAGAAATATTCATCATCTTCTTGGCTGCAATGATCAATATCTCT 5707
OY 698 ThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheValAsnGlyLysPro 717
Db 5708 CTATATGATATTCCTCTACACACACAGCTCTATCTTGACATAT---CAAGGGAAGAAAT 5764
OY 718 SerGlnGlnAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuVal 737
Db 5765 ATTTCTACTGCTAGTAAAGTGGTGGTTTACTTTGATTTCACTCA-----GACAGTCTGAG 5818
OY 738 AsnProMetArgGlyTyrArgGlnArgTyrSerIleGlnValGlySerSerGlyLeuVal 757
Db 5819 AGACCACTTAAGGGCATTTGCAAAACATAATTGTAATATGAGAGTGGTGGTCTTGGT 5878
OY 758 SerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrTyrSerPheGlyAsp 777
Db 5879 GGTAATGATTAAGCTTCTCAAGCAATGCTGCAAGATTTTACTCAAT----- 5932
OY 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyIleGlnAlaGlyTyrIle 797
Db 5933 -----TCAGAAGGTAAAGCAATATATATGATTTGGCGTACAGCTGAGGTGCAAGT 5983
OY 798 TrpSerAspAsnProPheAsnHisValProTyrArgLeuArgPhePheAlaGlyLysAspGln 817
Db 5984 TATAAGCAATAGTAAAAACCTGCTCCAGTATTTGACCGATTTTATTTATGGTGTATAGAT 6043
OY 818 SerIleArgGlyTyrAlaHisAspSerLeuSerProLleSerAspLys---GlyTyrLeu 836
Db 6044 AGTATAGAGATATGATACAGAAAGATCTTGACCAAAAGATCCTGCTTGGAGATGAA 6103
OY 837 ThrGlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGluPheMetLysAsp 856
Db 6104 ATTTGGTGTATAGAGATGGCTTTTCTTAACCTAGAGATATTTTGGACATTTCCAGCCAGAG 6163
OY 857 LeuArgLeuAlaVal-----PheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThr 874
Db 6164 CTAGCTCTTCATTAAGTTCATCTGATAGATAGATTCACAAAGATTTCTGTACAAACT 6223
OY 875 AsnAsp-----ThrLysIleGlyAlaGlyValGlyValAlaGlyTyrPheLysPro 890
Db 6224 TCTAACCAATCTCTTAAGTCAAAAGCAATCATATGGCTTGAACCTTGGTGGCTTACCA 6283
OY 891 ValGlyGlnValArg 895

```

```

Db 6284 ATGGAGATTTGGCA 6298
RESULT 18
US-09-689-065B-1
; Sequence 1, Application US/09689065B
; GENERAL INFORMATION:
; APPLICANT: Pfizer Products, Inc.
; TITLE OF INVENTION: LAWSONIA INTRACELLULARIS PROTEINS AND RELATED METHODS AND
; TITLE OR INVENTION: MATERIALS
; FILE REFERENCE: 3153, 00187/PC10589A
; CURRENT APPLICATION NUMBER: US/09/689, 065B
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: US Prov. 60/160,922
; PRIOR FILING DATE: 1999-10-22
; PRIOR APPLICATION NUMBER: US Prov. 60/163,858
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 6617
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-09-689-065B-1

Alignment Scores:
Pred. No.: 7,85e-09 Length: 6617
Score: 238.50 Matches: 198
Percent Similarity: 34.72% Conservative: 137
Best Local Similarity: 20.52% Mismatches: 391
Query Match: 5.05% Indels: 239
DB: caps: 38

US-09-914-168-2 (1-919) x US-09-689-065B-1 (1-6617)
OY 50 AsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGlnGlnIleGlnAlaArg 69
Db 3764 ATTTGCTGCTTCAAAAGACAGATCTTCTATTTGGTCTCCATTTCAATATTAAGCTCA 3823
OY 70 LeuAsnAlaAlaGlyLysAsnAlaLys---ProGlnSerGlnAlaLeuAspValAlaAsn 88
Db 3824 TCAAATGATGAGAGTTCACAAACAGAACATGCTTCTTGCAACGATTAAGAAAT 3883
OY 89 PheAspAspGlnSerProLleSerArgIleGlyGlnGlnSerProProLeuGlyLeuAsp 108
Db 3884 -----AAGGGA 3889
OY 109 MetSerValIleGlnGluThrThrProLeuSerLeuGlnGlnLeuPheAlaGlnGlnSer 128
Db 3890 TTTGCTGCTATCCCTAATAATCTGCATTAAATCTT-----CTATATTAACCAAAATATC 3943
OY 129 ThrGlnMetGlyLysAsnProAsnAspTyrIleProGluTyrGlnGlnGlnProAsn 148
Db 3944 TCCCACTTAATATTTCTACTGCAAAAAAGTA-----GCTCAACAACATCCAT 3991
OY 149 SerGlnValAlaValProProThrLeuGluProGlnLysProGlyLeuIleLysArgLeu 168
Db 3992 GCTGACTATGACTA----- 4006
OY 169 TyrAlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyr 188
Db 4007 TACGGCAGTTTCANATCAACACGCTGAAGAT-----TTTACT 4042
OY 189 GlnSerSerGlnSerGlyGlnThrSerAlaIleGlySerSer-----HisGlnLys 205
Db 4043 ATTTGATATAGTGGCTTATGATAGTACAGTGTGATGCTGACGTCATATATACATAGAA 4102
OY 206 ThrGluProTyrAlaAsnIleLysAlaAlaLeuGlnAspIleThrGlnGlnGlnSerAlaMet 225
Db 4103 AAACCAAAATTTATAGTAAATATTTGCTGTAACAGAACTTCTGTAACGTAATAGT--- 4159
OY 226 AspLeuAsnGlySerIleProArgLeuArgGlnThrValLeuValAlaAlaArgAlaVal 245

```

Db 4160 -----AATGGCTTATAAAGAA---AACACTATTGCTGATGTACGATTCAGGGCTT 4210
 Oy 246 GATYTYTAspIleaspleuSerIleIleArgAsnSerIleGlyValAlaSpvalIle 265
 Db 4211 AAGATTCTTGATCTCGATGATTAATCTTACACGACTGACTATTATTAAGGAGATCTACT 4270
 Oy 266 ILeHis-----AspleuGlyValProValTyr 274
 Db 4271 GATCATGCCCAAAATTAATGCAGAAATCAAAAATATGGGAATTAGA-----TATTTT 4324
 Oy 275 ILeAspTYTArgAla---ValGlyValArgGlyGlyAlaAspAspIleSerIle 293
 Db 4325 AGTGAATGCTCTGCAGATTTAGAAAGAACGGGGAAGA-----GCATTACTTGA 4375
 Oy 294 ThrValAlaAspGlyValProleuLeuIleGlyAspValPheHisIleGlyTyrGly 313
 Db 4376 TTTACTGTCAGAAAGAAAGCTTAAATTAACAGATGTTGTTCTCAAGGCTCAAAAGCTGA 4435
 Oy 314 ThrLysLysAsnLeuIleGlyAsnAlaSerAlaGlyHisGlyTyrPheAspGlyArgTyr 333
 Db 4436 AGTATGATTAACATCTCTGCTGCAATGAGTCTAAAAAGGA----- 4477
 Oy 334 LeuAspArgSerVal-----AspValIleLeuProAspAsnThrAlaSpValSerLeu 351
 Db 4478 -----TCAGTTATTAGTATGATGACTATGCTCCCAAGATATCCAAAAATTACCGAG 4528
 Oy 352 ILeTYTAspThrGlyThrGlnTyrArgPheAspGlyValValPhePheThrIleAspPro 371
 Db 4529 CTCTATGAAAAGAAAGCGCTACTATCTCGCTGA-----GTTAATTATGAAATTAAGAG 4582
 Oy 372 LysThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGlyLeuLeuGln 391
 Db 4583 AAAGAAATACCTCTCTGCA-----ACA 4606
 Oy 392 GlnLeuLeuThrValAlaAsnMetGlyGlyAlaTyrAsnLeuGlnAlaValArgAlaLeuSer 411
 Db 4607 CTATTGTTTACAGTAATAAGCGAAAACTTATATTAAAGATTCGCAATTGAAGA 4666
 Oy 412 AsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGlyLeuValPheProGln 431
 Db 4667 CTGGAACAATGAAAGCTAAA-----ACTTTAAAAAAGAGTTACCATTAACAGAA 4717
 Oy 432 Arg-----GlnGlnIleGlnAsn 437
 Db 4718 CATAATTTTTTATCATGCTTACTGGAACAGGTGATTACGTGAAGATATCTTGAACGT 4777
 Oy 438 AspGlnValSerPheGlnLeuInserSerSerArgThrGlnProAlaGlnValArgGly 457
 Db 4778 GACTCTATAGCAATC-----TCTGCCATATGCCATGAATCATGGCTATGATATTT 4828
 Oy 458 SerThrLeuGlnPro-----ValIleGlnThrValGlnLeu 469
 Db 4829 CAAGTTCCTACCTCACTAATCATTCATGAAAAAGAAATTTGTTTATTCATTTAAGTA 4888
 Oy 470 ThrAspGlyIleLeuMetAspIleSerProIleGlyPheSerAlaSerAlaSerIleGln 489
 Db 4889 AAAGAGAGTAACCCCTATAAATAGCAAAATAGACTTTAAAGAGATCTTATTGAGACA 4948
 Oy 490 AspLysLeuAsnLeuValAlaLysAlaArgHisLeuTyrAspMetProAspArg 509
 Db 4949 AATGAACAACCTCTTAACATAACAAAATTTGATGATCAATAAACATAGACAGATATTTT 5008
 Oy 510 ValLeuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerArg 529
 Db 5009 TCTCTTCTTGATGCAACATGAT-----GTAAAGCATTAACAGATTTTATTTACAGAT 5062
 Oy 530 AlaValSerAlaValAlaArgAlaIleLeuProAspGlySerGlnGln----- 546
 Db 5063 TATGCTTATGCAATTGCTGAATAGATCTT-----GAAACAACCAAAATATGAAGACATGCA 5119
 Oy 547 -----ValIleAspLeuProGlnArgThrAlaLeu----- 556
 Db 5120 ACAATTGATGTACTTCTTCTTATGATTAACAAAAGCTTCTTCTGTAAGATATTT 5179

Oy 557 -----AlaAsnArgLysThrProAlaSpValTyr----- 566
 Db 5180 GTTGAAGAAATACGTGTAAGATTAATGTTATCTCCGTGAATTAAGCCCTGCTGAT 5239
 Oy 566 ----- 566
 Db 5240 GCAGATCTTTTAAATGTCACATCTCCGACGCTTAATGAATGCTTAACGCTTGGC 5299
 Oy 567 -----GlnSerLysValProleu 573
 Db 5300 TATTTTAACAGTACATACAGATACAGTCCCTACAGGAAAGAAATGATGATGATCA 5359
 Oy 574 TyrValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrGlySer 593
 Db 5360 CTGTGAAGGT-----CAACAGCTCCAGACAGGTGCATCAACAGGTGCTGTGCTTAC 5413
 Oy 594 AspThrGlyThrArgLeuValThrLys-----PheGlnHisAsnLeuIleAsnArg 610
 Db 5414 TCACACATTTCAATTTGTTGCTTTCAGGAAGTATCTCAAGAAAGAACTTATGGGAAAA 5473
 Oy 611 AspGlyTyrGlnAlaGlyValGlyLeuArgLeuSerGlnAspLysGlyValLysLeu 630
 Db 5474 ---GGTTAATTTTAAAGTATGTAAGCTTTATTTTACGTAAGTCATCTTCTGATCTT 5530
 Oy 631 TyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuValArgAlaThrLeuGlyTyr 650
 Db 5531 TCTTTTACCAATCTCTGCTTTATGATACAGAC----- 5563
 Oy 651 GlnGlnGlyValPheGlyHisSerThrAsnGlyPheAspLeu----- 664
 Db 5564 -----TTTGGCTTATGATTAACATTTATACCTTACAGATGAATGGATGAC 5611
 Oy 665 -----SerThrArgThrLeuGlnHisGlyLeuSerArgSerIle 677
 Db 5612 TTCCGTAAGAAACCTTATGAGATACCATACGTATTTACCCCTATA----- 5659
 Oy 678 ILeGlnGlyGlyTyrPheAsnArgThrTyrSerLeuArgTyrArgLeuAspLysLys 697
 Db 5660 -----GGAAATATTCATCTATCTTGTGGCTATGAGATTCATGATTCGT 5707
 Oy 698 ThrGlnAlaProProGlnThrTyrPheAspLeuProValAspPheValAsnGlyLysPro 717
 Db 5708 CTATATGATATTCATCTCATCTGACGCCACGCTCTATCTTGCATAT---CAAGCGAAAT 5764
 Oy 718 SerGlnGlyAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeuVal 737
 Db 5765 ATTTCTAGTGTAGTAAAGTGTGCTTTTACTTTGATTTACA-----GACAGTCTGAG 5818
 Oy 738 AsnProMetArgGlyTyrArgGlnArgTyrSerLeuGlyValGlySerSerGlyLeuVal 757
 Db 5819 AGACCATCTAAAGGCGATATTGCAAAACATATGTTGATATGAGAGTGTGCTTTGGT 5878
 Oy 758 SerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp 777
 Db 5879 GGTAAATGATTAACCTTCTCAAGCCCAATTGTGTAACCAAGATTTTACTCAATT----- 5932
 Oy 778 AsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGlnIleGlyTyrIle 797
 Db 5933 -----TCAAGAGTAAGAAACCATTAATATCATTTGCGGTACACGTGACGCTGACGCT 5983
 Oy 798 TrpSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyIleAspGln 817
 Db 5984 TATTAAGATATGTAAGAAACCTGTGCGCAGATTTTGCACGATTTTATTTGCTGATTAAGAT 6043
 Oy 818 SerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLys---GlyTyrLeu 836
 Db 6044 AGTATTAAGAGATATGATACAGAGATCTTGGACCAACAAAGATCTCGCTGAGATGGA 6103
 Oy 837 ThrGlyGlyGlnValLeuValAlaGlyThrAlaGlyTyrAsnGlyGlyPheMetLysAsp 856
 Db 6104 ATTTGGTGTGATAGATGAGCTTTTCTTAACCTAAGATTAATTTTGGACATTTCAAGCCAGAG 6163


```

      Oy      857  LeuAtrGleuAlaVal-----PheGlyAspIleGlyAsnAlaTyrAspLysGlyPheThr 874
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      6164 CTAGGTCTTCATTAGTTCATTCATGACATAGATTCACAAACAGATTCTGTACAAACT 6223
      Oy      875  AsnAsp-----ThrLysIleGlyValGlyValArgTTrpAlaSerPro 890
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      6224 TCTAACCCATTCTCTAAACTCAACATCATATATGGCTTGAACCTTCGCTGGCTTACCA 6283
      Oy      891  ValGlyGlnValArg 895
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      6284 ATGGAGATTTCGCA 6298

RESULT 19
US-09-988-067B-7
: Sequence 7, Application US/09988067B
: GENERAL INFORMATION:
: APPLICANT: Haas, Rainer
: APPLICANT: Kleinhous, Harold
: APPLICANT: Tomb, Jean-Francois
: APPLICANT: Miller, Charles
: APPLICANT: Al-Garawi, Amal
: APPLICANT: Odenbreit, Stefan
: APPLICANT: Meyer, Thomas
: TITLE OF INVENTION: Helicobacter Polypeptides and
: TITLE OF INVENTION: Corresponding Polynucleotide Molecules
: FILE REFERENCE: 06132/040002
: CURRENT APPLICATION NUMBER: US/09/988,067B
: PRIOR FILING DATE: 2003-01-31
: PRIOR APPLICATION NUMBER: US 08/931,309
: NUMBER OF SEQ ID NOS: 112
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 7
: LENGTH: 3012
: TYPE: DNA
: ORGANISM: Helicobacter pylori
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (142)...(2682)
US-09-988-067B-7

Alignment Scores:
Pred. No.: 1,17e-06 Length: 3012
Score: 206.50 Matches: 202
Percent Similarity: 32.98% Conservative: 143
Best Local Similarity: 19.31% Mismatches: 318
Query Match: 4.37% Indels: 383
Gaps: 51
DB:

US-09-914-168-2 (1-919) x US-09-988-067B-7 (1-3012)
Oy      24  MetThrSerGlnAlaLeuAlaGlnGlnAsnAspProAlaAsnIleIleAsnHisValPro 43
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      25  TTGACTTCTCCAAAGAGAGCTCTCAAGATCTCAAAA-----AATGAAGCTCCA 75
      Oy      44  AlaHisAspThrAlaIleAsnGlnAlaLysAlaLysProProValLeuLeuThrPro 63
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      76  AAAAATGAGAGTTCAAGAGCTCAAAAGAAAGAAACCCCAATCCATCAACAGCT 135
      Oy      64  GlnGlnIleGlnAlaArg-----LeuAsnAlaIleGlyLeuAsnAlaLysProGlnSerGln 82
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      136  AAAAGAAAGAAAGTCAGTCCATTCTTATGTCGGGCTTTTTCATGCTGCACATGCTC 195
      Oy      83  AlaLeuAspValAlaAsnPheAspAspGlnSerProIleSerArgIleGlyGlnGlnSer 102
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      196  GCTATATGCAATTGTAAAGATT-----CGTGTGGCGGATATTGTG 234
      Oy      103  ProProLeuGlyLeuAspMetSerValIleGlnGlnUthrThrProLeuSerLeuGlnGln 122
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      235  GATTCTAAATAAATAGACACCGCTGCTTTTG-----GCT 267
      Oy      123  LeuPheAlaGlnGlnUthrThrGlnMetGlyLeuAsnProAsnAspTyrIleProGlnUthr 142
      |||      |||      |||      |||      |||      |||      |||      |||      |||

      Db      268  TTGTTCATCAA----- 279
      Oy      143  GlnGlyGlnGlnProAsnSerGlnValValProProThrLeuGlnProGlnLysPro 162
      Db      279  ----- 279
      Oy      163  GlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGlyValAsnLysValProArg 182
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      280  GGTATTTTAAAGACGTTTATGCC-----ACTTTGAAGCGGCATTA----- 321
      Oy      183  LeuLysAlaLysPheTyrGlnSerSerGlnSerGly-----GlnThrSerAlaIleGly 200
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      322  TTAGAGTTTCATTTTGATGATAAAAGCCAGATTGCCCGGGGTAGAAATTAACAGGTTATCGG 381
      Oy      201  SerSerHisGlnLysThrGlnProTyrAlaAsnIleLysAlaAlaLeuGlnAspIleThr 220
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      382  ACTGAAAGAAAGAAAGAGCGCTTAATCCAAATGGGAGTCAAAAGCGACACTTT 441
      Oy      221  GlnGlnSerAlaMetAspLeuAsnGlnSerIleProAlaGlyArgGlnThrAlaLeu--- 239
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      442  CATGAGCAAAATTTAGAG-----CATGCTAAACGGCTTTAAAA 480
      Oy      240  ValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIleIleArgAsnSerIle 259
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      481  ACCGCTTTAGAGCGGCGAGGCTATAT----- 507
      Oy      260  GlyGlnValAspValIleIleHisAspLeuGlyGlnProValTyrIleAspTyrArgAla 279
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      508  GGGAGCGTG----- 516
      Oy      280  ValGlnValArgGlyGlnGlyAlaAspAspLysAlaPheThrThrValAlaAsp----- 297
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      517  GTGAGAGTCCGACAGAAAGGTCAGTGAAGGCTATTTATGTCGCTGTGATGATGAT 576
      Oy      298  ---GluValProLeuLeuIleGlyAspValPheHisGlyLysTyrGlnUthrLysLys 316
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      577  AGCGGAGATGATTTATTTATCAAAACATCCATTATAGGGAAGCGGAAATTAAGCC 636
      Oy      317  AsnLeuIleGlnAsnAlaSerIleGlnHis-----Gly 327
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      637  CGCATATTTGAATCTTTGATGGCAACAGACAGATTTTCATGGGCTGGATGTGGGGC 696
      Oy      328  TyrPheAspGlyArg----- 332
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      697  TTGAATGAGCGGAATTTGCTTATAGATCAACTAGATTCATTCGCTATCCAAAGAT 756
      Oy      333  -----TrpLeuAspArgSerValAsp-----ValIleLeuProAsp 344
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      757  GTGTATATGCGTAGGGCTTACTTAGACGCTCATATTTCTTCCGCTTTTGTGAAGAGGAT 816
      Oy      345  ---AsnThrAlaAspValSerLeuIleTyrAspThr-----GlyThrGlnTyrArgPhe 361
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      817  TTTTCTACCATGACGCTTAAGCTTATTAAGTCAAAAGAGGGATCCAAATACAGATT 876
      Oy      362  AspGlnValValPhePheThrIleAspProLysThrAsnGlnLeuUthrThrAspProAsp 381
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      877  TCAACACTTTTA-----ATAGAGATTGGCAACCCG 906
      Oy      382  LysLeuProValLysArgGlnLeuLeuGlnGlnLeuLeuThrValAsnMetGlyGlnAla 401
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      907  GTAGTCCCTTAAAA-----ACCTTAGAAAGACCGCTTAAGTGAAGAAAGAAAGATGTC 960
      Oy      402  TyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeu-----IleAlaThrArg 418
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      961  TTTAATATTTAGCATTTAAGAGCGGATTTTAAAGAACCGAATTCGCGCATAC 1020
      Oy      419  TyrPheAsnMetValAsnThrGlnIleValPheProGlnUthrGlnGlnIleGlnAsnAsp 438
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      1021  GGTATAGCTTTGCG-----GTGGTAGAGCCAGACTTGATTA----- 1059
      Oy      439  GlnValSerPheGlnGlnSerSerSerSerArgThrGlnProAlaGlnValAspGlnSer 458
      |||      |||      |||      |||      |||      |||      |||      |||      |||
      Db      1060  -----GATGAAGAAA 1068
```


Oy	459	Thr---	LeuGIuProValIleGIuThValGIuIleuThAspGlyIleIleuMetAsp---	476
Db	1069	AACGGGCTTGGTGAAGTCAATATTATCGTAAATGAGTGGCGATATGGGTGATATACATGAT	1128	
Oy	477	-----IleSerProIleGIuPheSerAlaSerAsnIleuIle-----	488	
Db	1129	GTCAATCATTTTACAGGAACCCAGCAGAGGATAGATCATCTTGAAGGGAGATTATTGTTA	1188	
Oy	489	-----GIuAsnArgLysIleAsnIleValAlaAlaLysAlaArgHisIleuTyArgMetPro	506	
Db	1189	GGGCGTAAAGGATAAATACAACTTGATCCAAACTGAGA-----	1224	
Oy	507	AspAspArgValIleuAlaIleAsnHisAspGlyValAlaAspSerIleIleuGIuArg	526	
Db	1225	-----AAATCCGAAATTTCTTAAGCGCT-----TTAGCATTC	1257	
Oy	527	IleSerAspAlaValSerAlaValAlaArgAlaIleIleuProAspGluSerGIuAsnGIu	546	
Db	1258	TTCTCTAAAGTCAAAATTTGAAGAAAAAGGGTT-----AATAGCTCA	1299	
Oy	547	ValIleAspIleuProGIuArgThAlaIleuAlaAsnArgLysThrProAlaSpValIlyr	566	
Db	1300	CTCATGGATTAA-----	1311	
Oy	567	GIuSerLysValProIleuTyValPheValAlaSerAspLysProArgSpGIuGIu	586	
Db	1312	-----TTAGTGAGCGGTAGAGAGGGCGGTACTGGCGAG	1344	
Oy	587	Ile-----GIuIleuGIuTyProGIuLysSerAspThrGIuThArgIleuValThIlySpheGIu	604	
Db	1345	TTTGCATTTTGGGTAGGCTATGAGCTCTTAAGAGCGCTTATGCTT-----AAT	1392	
Oy	605	HisAsnIleuIleAsnArgAspGlyTyArgAlaGIuAlaGIuIleuArgLysIleuSerGIuAsp	624	
Db	1393	GGGAGCGTGAGCCGAAGAACTTTTGGCCAGAGGCGAAGCATGAGCTTGATCTGATAC	1455	
Oy	625	-----LysLysGIuValLysLysIleuTyAlaThrLysPro-----	635	
Db	1453	ATCGCTACAGGGGGGTAGTCTTATCCGGGATGCCAAAGAGCGGGCGATGTTT	1512	
Oy	636	-----LysSerHisProIleuAsnAspGIuIleuArgAlaThrIleuGIuTy	650	
Db	1513	GCCGGGAATTGGACCTTGACATACCA-----	1538	
Oy	651	GIuGIuGIuValPheGIuHisSerThrAsnGIuPheAspLysIleuThrArgThIleu---	669	
Db	1540	-----AGGATTTT-----GACAGCTGGTATAGCTCTACATCAACTTAT	1581	
Oy	670	---GIuHisGIuIleSerArgSerIleIleGIuAsnGIuGIu---	682	
Db	1582	GCGGATTACAGGAATACCTCAATATCAACAAGCGGGCGCTTGGGTGAATGTC	1641	
Oy	683	-----TrpAsnArgThrTy---SerIleuArgTyArgIleuAspLysLysLys	697	
Db	1642	GGGGCATCTCGGTATATAGAACCACTGTGAGCTTGAGGTATACCTGAACTTTACCAA	1701	
Oy	698	ThrGIuAlaProProGIuThrTrpGIuAspLysIleuProValAspPheValAsnGIu-----	715	
Db	1702	CTCTTGTTGTTACAGCACCCCTTTATACAAACCGCTACTATCTCTGTTAATGAACGTG	1761	
Oy	716	LysProSerGIuGIuAlaIleuAlaGIuValAlaValHisLysTyThrValAla-----	733	
Db	1762	TTCTCCAAAGCAATGTCTACCCCCGATCGGTATTTATACATGCGCTTATACAGCGGTAA	1821	
Oy	734	-----AspAsnIleuValAsnPro-----Met	740	
Db	1822	AACCCCTTACAACTGAAACGTGTCTAGTCGTGAGAGCATCACCACTTACACAGAAATA	1881	
Oy	741	ArgGIuTyIleArgIleu-----TyrSerIleuGIuVal	751	
Db	1882	AGAGGATTTTGGGATAGGATATACATACGCTATACCAAGAGCTCTTACACCTTGATATG	1941	

```

OY 752 GLY-----SerSerGlyIueValSerAspIasn 761
Db 1942 AGCATATGACAAACACCCGATGATATTACTTCCTCCAGAAATGGGGTTATCTTTAGT----- 19959
OY 762 MetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAsp----- 777
Db 1996 -----TCCTATGGACGAGATGTCTGGCTTCCAAAGCCTTCGACGCTCAATTCTTGG 20466
OY 778 AsnAlaTyrGlySerAsnAlaGlnHisGlnMetThrGlyGly----- 791
Db 2047 AACGGGTTAGCGGGGAATGTCCTGACACCAAGTTTATGTGAATTTCGCCCTTACCAC 21066
OY 792 -----IleGlnAlaGlyTyr 796
Db 2107 CATTTGCAAAATATTTATTATGATAGATTTGATCGCTCGCTTTAAACGCAAGAGATAT 21666
OY 797 IleTyrSerAspAsnPhe---AsnHisValProTyrArgLeuArgPheAlaGly 815
Db 2167 ATCTTTAGATATAACCCGATGATCTTCCTGACCTTAACCTCCACCTTCTACATCGGGGCG 22266
OY 816 AspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspGlyTyr 835
Db 2227 GTAACCAACGCGAGAGCCTTTAGGAACGAGATCGGTACTCTCAAGATGAGTTTGCTTG 22866
OY 836 LeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyrGlnPheMetLys 855
Db 2287 TGGCTTGAGAGCGCATGGGATTTTACCCTGCTTACAGATTAAGCTATGGGTGCTAAG 23466
OY 856 -----AspLeuArgIleuAlaValPheGlyAspIleGly----- 866
Db 2347 GCGGCTAAATAGCGCTAGCGTGGTGTTTTGACTTGTGGTTCTTAACCTTTAAACCCCA 24066
OY 867 -----Asn 867
Db 2407 ACTAGAGGAGATTTTCTCTAATACGCTCCGTTAGCACAGCGAATTTTAAGATTAATGCG 24666
OY 868 AlaTyrAspLysGlyPheThrAsnAspThr---LysIleGlyAlaGlyValGlyValArg 886
Db 2467 GTTATTAGGGCGCTGGGTTTGAAGAAGCAGCTGGAGGGCTTCCACAGCGCTTGCAATGTGA 25266
OY 887 TrpAlaSerProValGly 892
Db 2527 TGGATTTCCGCCCATGGG 2544

RESULT 20
US-10-335-977-157
Sequence 157, Application US/10335977
GENERAL INFORMATION:
APPLICANT: DOUGLAS SMITH et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESS: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD/ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
Prior APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-Dec-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.

```



```

Db 1405 -----TTA 1407
Oy 577 ValAlaSerAspLysProArgAspGlyIle-----GlyLeuGlyIrrProLysSerAsp 594
Db 1408 GAGACCGTAGAAGAGAGGCGACCGCGAGTTGCAATTCGAGTTGGCTATGGCTCTTAT 1467
Oy 595 ThrGlyThrArgLeuValThrLysPheGluHisLeuIleAsnArgAspGlyTyrGln 614
Db 1468 GGAGGCGTCATGCTT-----AATGGAGCGTAGGAGGAAGAATCTTTTGGC 1515
Oy 615 AlaGlyAlaGluLeuArgLeuSerGluAsp-----LysLysGlyValLysLeuTyrAla 632
Db 1516 ACAGGGCAAGAGCATGAGCTTGTATGCTACATTCGCACAGCGGCGGATCTTATCCG 1575
Oy 633 ThrLysPro-----LeuSerHisProLeu 640
Db 1576 GGCATGCCAAAAGGGGGCGGCTATGTTGCCGGGAATTGGAGCTTGACTATATCCA--- 1632
Oy 641 AsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGlnValPheGlyHisSerThrAsn 660
Db 1633 -----AGGATTTT-----GAC 1644
Oy 661 GlyPheAspLeuSerThrArgThrLeu-----GluHisGluIleSerArgSerIleIle 678
Db 1645 ACCTGGTATAGCTCTACGATCAATCTTTATGGGATTACAGAGATTAAGCATCCATACATC 1704
Oy 679 GlnAsnGlyGly-----TrpAsnArgThrTyr 687
Db 1705 CAACAAGGGGGGGCTTTGGGGAATGTCGGCGCATCTGGGATATAGAACCCATGTG 1764
Oy 688 SerLeuArgTyrArgLeuAsp-----LysLeu-LysThrGlnAlaProProGluThr 705
Db 1765 ACCTTAGGGTATACCTGATGTAATGTACCAAACTCTGTGGTTGACGAGCCCTATACAA 1824
Oy 705 PclnAspLeuProValAspPheValAsnGlyLysProSerGlnGlnAlaLeuAlaG 725
Db 1825 GCCTACTATCTCTGTATATGAATG--GCCTCTCCAAAGCAATGTTCACACCCGCAAC 1883
Oy 725 ValAlaValHisLys-----ThrValAlaAspAsnLeuValAs 738
Db 1884 GGTGATTATACACCGCTTATCAGGGGTAGACTCCATTTGGTCCGAAAGCGTTCTAG 1943
Oy 738 nPro-----MetArgGlyTyrArgGlnArg----- 746
Db 1944 TCCTGAGGAGATCACCACATTCACAGAAATAAAGSTATTGGATGAGGATTACACAC 2003
Oy 747 -----TyrSerLeuGlnValGly----- 752
Db 2004 GCCTATCACCACTTCTTACCCCTTGATGTAGCTATGACACACCGCATATTATTAT 2063
Oy 753 ----SerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerG 771
Db 2064 CCCTAGAAATGGGCTTATCTTAGT-----TCTATGCGACAATGCTCTGG 2108
Oy 771 ValTyrSerPheGlyAsp-----AsnAlaTyrGlySerAsnArgAlaHisG 787
Db 2109 TTGGCAACGCTCGCACCTCAATCTTGAAACGGGTTAGGGGAATGTCGTTAACAC 2168
Oy 787 nMetThrGlyGly----- 791
Db 2169 CAAGATTATGGTAATATTCGCGCTTACACCATTTGCAAAATATTATTGATAGATT 2228
Oy 792 -----IleGlnAlaGlyTyrIleTyrSerAspAsnPhe-----AsnHisVa 805
Db 2229 GATCGCTGCTTTTAAACCAAGGCGGTATATCTTAGTATACACCGATCATTTACT 2288
Oy 805 LProTyrArgLeuArgPhePheAlaGlyLysAspGlnSerIleArgGlyTyrAlaHisAs 825
Db 2289 GCCCTTAAATCACACTTCTACATGGGGGCGTAACACGAGTGAAGGCTTTAGCAACG 2348
Oy 825 pSerLeuSerProIleSerAspLysGlyTyrLeuThrGlyGlnGlnValLeuAlaValG 845
Db 2349 CTCATATCACACCTAAGATAGTTGGCTTGGCTTGGAGGATGGGATTTTACCGC 2408

```

```

Oy 845 yThrAlaGluTyrAsnTyrGluPheMetLys-----AspLeuArgLeuAlaValPheG 863
Db 2409 TTCCTACTGATTAAGTACTATAGGGGTGTAAAGCGGCTTAAATGCGTTAGCGGTTT 2468
Oy 863 yAspIleGly----- 866
Db 2469 TGACCTTGGTTTCTTAACCTTAAACCCCACTAGAGGGGATTTCTCTATTAACGCTC 2528
Oy 867 -----AsnAlaTyrAspLysGlyPheThrAsnAspH 877
Db 2529 CACACAGAGCGGAATTTAAAGATTATAGCGCTTGAGGGCGGTGTGAAAGGGCGAC 2588
Oy 877 r---LysIleGlyAlaGlyValGlyValArgTyrPheAlaSerProValGly 892
Db 2589 TTGGAGGCTTCTACAGGCTTACAGATTGAATGATTTGCCCATGGG 2637

RESULT 21
US-10-419-128-4442
; Sequence 4442, Application US/10419128
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/10/419,128
; PRIOR APPLICATION NUMBER: US/09/252,991
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4442
; LENGTH: 2466
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-419-128-4442

Alignment Scores:
Pred. No.: 4.3e-05 Length: 2466
Score: 186.00 Matches: 175
Percent Similarity: 33.74% Conservative: 132
Best Local Similarity: 19.23% Mismatches: 276
Query Match: 3.93% Indels: 330
DB: 8 Gaps: 42

US-09-914-168-2 (1-919) x US-10-419-128-4442 (1-2466)
Oy 223 SerAlaMetAspLeuAsnGlySerIlePro-ArgLeuArgGlnThrAlaLeuValAla 242
Db 5 TCTGTAGTATGATGTGAATTCGAATCTGCGCCCTTTGGCGGCAATCTCTATTGTCAGT 64
Oy 242 aArgAlaValAlaGlyTyrTyrAspPheLysPheSerIleIleArgAsnSerIleGlyL 262
Db 65 TGGAAAAGAGGAGCTCCAGAA---ACGCTTCTGCTACCCGCGGCTGACGCCCTGA 121
Oy 262 LAsp-ValIleIleHisAspLeuGlyLuproValIlyrIleAspTyrArgAlaValG 282
Db 122 TCAATCCCGAGGTTTAC-----GCCGAGTCCCTTCACTGTTCCGATA 163
Oy 282 aLArgGlyGlnGlyAlaAspLysAlaPheThrThrValAlaAspGlnValProLeu 302
Db 164 TCGGGTCAATGCGCTGCAAGCGGTGTCCGCGGTAGCGTTCGCCCGCTCCCTGA 223
Oy 302 euIleGlyAsp---ValPheHisHisGlyLysTyrGluThrLysLysAsnLeuIleG 321
Db 224 ACCTCGGAGACACATGACGACGACGAGCCCTGGTCCAGGCAACCCGCTCGTTCAAGA 283
Oy 321 snAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPheAspArgSerValAspVal 341
Db 284 CC-----GGCTTCTTCCAGGACATCCAGCTCGGCGTGTATGGCAACGTGC 328

```

341 leleuproaspasnthalaspaValSerleuIleTyrAspThrGlyThrGlnTyrArgp 361
 Db 329 TG----- 330
 Oy 361 heaspGluValAlaPhePheThrIleaspProLysThrAsnGlnLeuThrThrAspPro 381
 Db 331 -----GTACTCAGTACTGACGAGCCCTCCATCTCCAGCATTCAGATCGAGGGCA 382
 Oy 381 sPLysLeuProValIysArgGlnLeuGlnGlnLeuLeuThrValAsnMet-----G 399
 Db 389 ACAAGGCGATTTCAGAGCAAGCACTGCTCAAGGCGCTCAACAGTCCGCGCTGGCCGAGG 442
 Oy 399 lyGlnAlaTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeu----- 414
 Db 443 GCGAGATCTTCAGAGCGCCGACCTCGAAGGCGCGGACAGAGCTGCACAGCCGCACTAG 502
 Oy 415 --IleAlaThrArgTyrPheAsnMetValAsnThrGlnIleValPheProGlnArgGln 434
 Db 503 TGGCCGAGGCGCGCTACTGCGCGAGATCAACGCGAGATGCCGCGCGAGCCGAGACC 562
 Oy 434 IuIleGlnAsnAspGlnValSerPheGlnGlnSerSerSerArgThrGlnProAlaG 454
 Db 563 GC----- 564
 Oy 454 InValAspLuserThrLeuGlnProValIleGlnThrValGlnLeuThrAspGlyIle 474
 Db 565 -----GTCCGCTGAAGATCAACATCAACAGCAAGAGGACGG 598
 Oy 474 euMetAspLuserProIleGlnPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnL 494
 Db 599 TCGCCGCGATTTCACATCAACGCTGTGGCGAACGCTCTTCTCCGAGAA----- 651
 Oy 494 euValAlaIaLysAlaAlaGlnIleLeuTyrAspMetProAspArgValIleAla---I 513
 Db 652 -----GACSTGACCGGACCTGTGCACTGAAAGACCACTGGCTGTCTTCT 700
 Oy 513 leAsnHisAspAspGlyValAsnArgSerIleLeu----- 524
 Db 701 TCAGAAACGACGACAGTACGCGCGAAACCTCTCCGCGACCTCGAGCGCTCGCT 760
 Oy 525 -----GlyArgIleSerAspAlaValSerAlaValAlaArgAlaIle 539
 Db 761 CCTATTACTGACCGCGCTTACATCAACATGATATCGCTCCACCCAGGTATCTCA 820
 Oy 539 euProAspLuserGlnAsnGlnVal-----IleAspLeuProGlnArgThrAlaL 556
 Db 821 CCGCGGACAGAGACGCTTACATCACTCAACATCAACGAGGGTGAAGATACACCA 880
 Oy 556 euAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValPro----- 572
 Db 881 TCCGCGACGTGAAGCTGACCGCGGACCTG-----AAGGTCCCGAAGAGAGAG 928
 Oy 573 -----LeuTyrValPheValAlaSerAspLysProArgAspGlnIleGlyL 589
 Db 929 TGAAGCCCTGCTGCTGCTA-----CAAGGCGCGAGTGTTC-C 966
 Oy 589 euGlyTyrGlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleA 609
 Db 967 TCGCGCAAGCGTGAATGAC-----CACCACTCTC----- 992
 Oy 609 snArgAspGlyTyr-----GlnAlaGlyAlaGlnLeuArgLeuSerGlnAspL 625
 Db 993 --CGACCTAATCAACCCCTCGCTGCGTGAAGGAGGTACACCTTCGCGCAACCTTAAGCGC 1050
 Oy 625 yAlsGlyValIleValLeuThrAlaThrLysProLeuSerHisProLeuAsnAspLIne 645
 Db 1051 GTCCCGGAGAG--CCCATGACAGACAGCAAGACCGTCTCGTAGCT---TCTGCTCGAGACC 1106
 Oy 645 rGAlaThrLeuGlyTyrGlnGlnGlnValPheGlnHisSerThrAsnGlyPheAspLys 665
 Db 1107 GGGGACAGCGC--GCCTACGTCAACCGCATTC-----AATTTCGCGGCA 1147

Oy 665 eThrArgThrLeuGlnHisGlnIleSerArgSerIleIleGln--AsnGlyGlyTyrP 684
 Db 1148 ACACCAAGACCGAGACAGATGCTGCGCGCGAAATCGCGCAGATGAGAGCGGCTGGG 1207
 Oy 684 snArgThrTyrSerLeu-----ArgTyrArgLeuAspLysLeu----- 696
 Db 1208 CCTCCACTCTACTGTATCAGCAGCAGTCCAGAGGCGCGCTGAGCGCTCGCTACTTCAAG 1267
 Oy 697 -----LysThrGlnAlaProProGlnThrThrGlnAspLeuProValAspPhe- 712
 Db 1268 AAGTCAACCTGAGACACCGCGCGGTCTCCCGGACCGAGACAGGTGACAGTCAACTCA 1327
 Oy 713 --ValAsnGlyLysPProSerGlnGln-----AlaLeuLeuAlaG 725
 Db 1328 GCGTCAAGAGCAACCGTCCGCGCTGATACCGCCAGCGTGGCTTCCGCGAGCGCGC 1387
 Oy 725 lyValAlaValHisLysThrValAlaAspAsn----- 735
 Db 1388 GCGTATCTCTCGCGCGCTGATGACAGACAACTCTCTGGGTACCGGCAACAGTCA 1447
 Oy 736 -----LeuValAsnPro- 739
 Db 1448 GCATGCGCTGACCGGACGGAATACAGACCGCTACAACTTCGCTTCGTCACCCCT 1507
 Oy 740 -----MetArgGlyTyrArgGlnArgTyr----- 747
 Db 1508 ACTGACCGCTGACGCGCGCTGACCGCTGCTGCTACAGCGCTTCTCGGAGAGACCGACTAG 1567
 Oy 748 -----SerLeuGlnValGlySerSerGlyLeuValSerAspA 760
 Db 1568 ACGAACTCGACGTCGACGTCGCGCAGTACTCGGTGAAACGCTGGGC-----G 1615
 Oy 760 laAsnMetAlaIle----- 764
 Db 1616 CCGCATGAGCATCGGCTACCGCATCAGCAGACCTGCGGCTGAGCTTGGCTGAGCG 1675
 Oy 765 -----AlaArgAlaGlyIleSerGlyValTyrSerPhe----- 775
 Db 1676 TGCAGCGCACACGATGACACCGCGCTTACACCGTGAACAGATCTGACACTTCGCG 1735
 Oy 776 -----GlyAspAsn-----AlaTyrGly----- 781
 Db 1736 ACNAGAGAGCGCACACTTACCACTTCAGAGGCTTCATCGCGCTGCTGCGAATCGACC 1795
 Oy 782 -----SerAsnArgAlaHisGlnMetThrGlyGlyIleGlnAla---- 794
 Db 1796 TGAACAAGGCGTGTGGCCACCGCTGTACTTCCGACAGCTGACCTCGGAACACACC 1855
 Oy 795 -----GlyTyrIleTyr----- 798
 Db 1856 TGGCGGCGACGCACTGTGCTTCTACAGATCGACTACCGCGGCGCAAGTCTGCGCCGC 1915
 Oy 799 --SerAspAsnPhe----- 802
 Db 1916 TGACCGCACACTACCACTACCGCTTCCACACGAGCTGGGCTATGGCGACGCTACGGAT 1975
 Oy 803 -----AsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg 821
 Db 1976 CCAACGAGCGCTGCGCTTCTACGAGACACTACGCGCGGCTTCAACTCGGTGCGGG 2035
 Oy 821 lyTyrAlaHisAspSerSerProIleSer----- 831
 Db 2036 GCTTCAAGACAGACAGCGCTGCGGCTGTAGTACCCCTTAGGGCTTACAGAGATGGCA 2095
 Oy 832 -----AspLysGlyTyrIleThr----- 837
 Db 2096 AGATCATCTCTGGCGCTGATGAGGTGTGTGCTACACGAGCCGAGACAGATCCGAGAG 2155
 Oy 838 -----GlyGlyGlnValIleValAlaGlyThrAlaGln-----TyrAsnTyrGlnPhe 854
 Db 2156 CTTTCGCTGTGTAACATCTGTATGACGCGCGGCGAGAGTGTGTGCTGCGCGCTTGG 2215
 Oy 854 eLysAsp-----LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp- 870

```

Db 2216 TCAGAGCCAGCGCCAGTGGCTACCGTCTTCGGAGCGTGGTACACCTTCGATA 2275
Oy 871 -----LysylPheThrAsnAspThrLysIle----- 879
Db 2276 CCGACTGCCCGACCAAGACACCACCACTGGAGCGCATCTCAAGACCAACCTCGCCA 2335
Oy 880 --GlyAlaGlyValGlyAlaGlyAlaGlyAlaGlyAlaGlyAlaGlyAlaGlyAla 899
Db 2336 GTTCGGTGGCGGTGGCGCTGACCTGATCACCAGCGGTGGCGGTGACCTTCAGCTGG 2395
Oy 899 IaThGlyValLysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 907
Db 2396 CGACGCGCATCAAGAACCGCGGACAAC 2421

RESULT 22
US-10-366-683-4442
; Sequence 4442, Application US/10366683
; GENERAL INFORMATION:
; APPLICANT: Rubenfeld, Marc J.
; APPLICANT: Nolling, Jork
; APPLICANT: Delouphery, Craig
; APPLICANT: Bush, David
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: PAT03-04
; CURRENT APPLICATION NUMBER: US/10/366,683
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: 09/252,991
; PRIOR FILING DATE: 1999-02-18
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4442
; LENGTH: 2466
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-366-683-4442

Alignment Scores:
Pred. No.: 4,3e-05 Length: 2466
Score: 186.00 Matches: 175
Percent Similarity: 33.74% Conservative: 132
Best Local Similarity: 19.23% Mismatches: 276
Query Match: 3,938 Indels: 350
DB: 9 Gaps: 42

US-09-914-168-2 (1-919) x US-10-366-683-4442 (1-2466)
Oy 223 SerAlaMetAspLeuAsnGlySerIlePro-ArgLeuArgGlnThrAlaLeuValAlaI 242
Db 5 TCTGTAGTCATCTTGAATTCGATCGCCGCTTGGCGGAGATTCTCTATTGTCTCACT 64
Oy 242 aArgAlaValAlaGlyTyrTyrAspIleAspSerIleIleArgAsnSerIleGlyGluVal 262
Db 65 TGGAAAAAAGGACTCTCAGAA---ACGCTTCTGCTACCCGCGCTGTGACGGCTGA 121
Oy 262 IAsp-ValIleIleHisAspLeuGlyGluProValTyrIleAspTyrArgAlaValGluVal 282
Db 122 TGATCGCGAGGTTCAC-----GCCAGTCCTTCACTGTTCCGATA 163
Oy 282 aIArgGlyGlyGlyAlaAspAspLysAlaPheThrThrValAlaAspGluValProLeuL 302
Db 164 TCCGGGTCAATGCGCTGACGCGGTGTCCGCGGTAGCTGTTCGCCGCCCTGCGCTGA 223
Oy 302 euIleGlyAsp---ValPheHisIleGlyLysTyrGluThrLysLysAsnLeuIleGluVal 321
Db 224 ACGTGGCGAGACCATCGACGACGAGCCCTGCTCCAGGCCACCGCTCTGTTCAAGA 283
Oy 321 snAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrPheAspArgSerValAspValI 341
Db 284 CC-----GGCTCTTCCAGACATCAGCTCGCGCTGTATGGCAACGTGC 328
Oy 341 IeLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThrGlyThrGlnTyrArgP 361

```

```

Db 329 TG----- 330
Oy 361 heAspGluValAlaPhePheThrIleAspProLysThrAsnGlnLeuThrThrAspProA 381
Db 331 -----GTACTCACTAGTACGAGCCCGCTTCATCTCCAGCATCGATCGAGGCA 382
Oy 381 sPlyLysProValLysArgGlyLeuLeuGlyGlnLeuLeuThrValAsnMet-----G 399
Db 383 ACAAGCGCATTTCCAAAGAACCTGCTCAAGGCCCTGAACAGTCCGCCGTGGCGGAG 442
Oy 399 IyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeu----- 414
Db 443 GCGAGATCTTCCAGCGCGCCACCTCGAAGGGTGGCGGACGACGACCAACCCAGTACG 502
Oy 415 --IleAlaThrArgTyrPheAsnMetValAsnThrGluIleValIlePheProGluArgGluG 434
Db 503 TGGCCCGAGGCGCGCTACTCTGCGCGGAGATCAACCGGAAATGATCCGCGCGGCAACC 562
Oy 434 IInIleGlnAsnAspGlnValSerPheGlyGlnSerSerSerArgThrGluProAlaG 454
Db 563 GC----- 564
Oy 454 InValAspGluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleL 474
Db 565 -----GTGCGCTGAAGATCAACATCAACAGAGCGACGG 598
Oy 474 euMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLysLeuAsnL 494
Db 599 TCGCCCGCATTTCCACATCAACGTCGTGGGCAACCGCTCTTCCGAAGA----- 651
Oy 494 euValAlaAlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeuAla---I 513
Db 652 -----GACCTGACCGACCTGTGAACTGAACTGAAGACCAACCACTGGCTGTCT 700
Oy 513 IeAsnHisAspAspGlyValAsnArgSerIleLeu----- 524
Db 701 TCAGAGACGACGACAACTACGCCCGGAAAGCTTCGCGGACCTCGAGCGCTGCGCT 760
Oy 525 -----GlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleL 539
Db 761 CCTATTACCTGGACCGCGCTACATCAATGATGATTCGCTCCACCGCATTCATCA 820
Oy 539 euProAspGluSerGluAsnGluVal-----IleAspLeuProGluArgThrAlaL 556
Db 821 CCCCGCAAGAACGACGCTACATCACCCTCAACATCAACAGGGGTGAGATGACACA 880
Oy 556 euAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValPro----- 572
Db 881 TCCGCGACGTGAAGCTGACCGCGACCTG-----AAGGTCGCGAAGAGAGAG 928
Oy 573 -----LeuTyrValIlePheValAlaSerAspLysProArgAspGlyGlnIleGlyL 589
Db 929 TGAAGCGCTGCTGCTGTA-----CAGAAGGCCAGAGTTC-C 966
Oy 589 euGlyTyrPlySerAspThrGlyThrArgLeuValThrLysPheGlnHisAsnLeuIleA 609
Db 967 TCGCGCAAGCTGATGAC-----CACCAACTC----- 992
Oy 609 snArgAspGlyTyr-----GlnAlaGlyAlaGluLeuArgLeuSerGlyAspL 625
Db 993 --CGACCTAATCAACCGCTGCGCTGGGCAACGAGGGCTACACCTTCCCAACGCAAGCG 1050
Oy 625 ySlyGlyValAlaLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuA 645
Db 1051 GTGCGGGAAG-CCCATGACGACGACAGACCGCTTCGGTGACT--TGTGCTGCAACC 1106
Oy 645 rGAlaThrLeuGlyTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLys 665
Db 1107 GGGCAAGGCC-GCTACGTCACACCGCATC-----AATTTCGCGGGA 1147
Oy 665 erThrArgThrLeuGlnHisGlyIleSerArgSerIleIleGln---AsnGlyGlyTPA 684
Db 1148 ACACGAAGACGAGACGAAAGTGTGCGCGGAATTCGCGCAGATGAAAGCGCGCTGGG 1207

```

Oy	684	snaIghTyrSerLeu-----ArgTyrAlglEuAspLysLeu-----	696
Db	1208	CCTCCACCTACTCATGATGACCAGGTGCAGAAGGCCCGCTGGAGCGCCTGCTACTTCAGG	1267
Oy	697	-----LysThrglnAlaProProlThrTrpGlinspLeuProValAspPhe-	712
Db	1268	AAGTCACAGCTTGAAACCCCGCGGTTCCCGCACCGAACACAGCTGAGCTCAACTACA	1327
Oy	713	--ValAsnGlyLysProSerGlnGlu-----AlaLeuLeuIag	725
Db	1328	GCGTCGAGAGACCAACCCGCTCCGGCTGCATCACCGCCAGCTCGCTTCGCCAGAGCGCG	1387
Oy	725	LysAlaAlaValHisLysThrValAlaAspAsn-----	735
Db	1388	GCGTCGATCTCGCGCGGCTTCGATCGATCACCGAAMCAACTTCCTGGTAACGGGCAACAAGTCA	1447
Oy	736	-----LeuValAsnPro-	739
Db	1448	GCATCGGCGCTGACCCGCGAGCAATACCAAGCCGCTACCACTTCGGCTTCGTCGACCCCT	1507
Oy	740	MetArgLysTrpArgGlnArgTyr-----	747
Db	1508	ACTGGAACCTTCGACGCGCGTCAACCTTCGGCTGACACCGCTTCTACCGGAAGACCGACTACG	1567
Oy	748	-----SerLeuGlnValGlySerSerGlyLeuValSerAspa	760
Db	1568	ACGAACTCGACGTCGACGTCGCGGCTACTCGTGAAACAGCCTGGGC-----G	1615
Oy	760	IaAsnMetalIaIle-----	764
Db	1616	CCGCGCATGACATCGCGCTTACCCGATPCAGSAGAACCTCGCGCTGACTTATGGCCTPGAGCG	1675
Oy	765	-----AlaArgAlaGlyIleSerGlyValYalYrSerPhe---	775
Db	1676	TGCAAGCGCACAGATCGACACCGCGGCTGACACCGTAGACAGATCTACGACTTCCTCG	1735
Oy	776	-----GlyAspAsn-----AlaTrpGly-----	781
Db	1736	ACAAAGAAAGCGCAACTTCACCACTTCAGAGCTTCGATCGGCTGGTCCGAAATCGACCC	1795
Oy	782	-----SerAsnArgAlaHisGlnMetThrRelayLyIleglnAla----	794
Db	1796	TGAACAAGGGCGGTGCGCCACCGGTGCATCTCGAGAGACCTGACCTCGAAMAACACCC	1855
Oy	795	-----GlyTyrIleTrp-----	798
Db	1856	TGCGGGGACAGCACTGTGCTTCTACAGATCGACTACCGGCGCAGGTTCGCCCGCG	1915
Oy	799	--SerAspAsnPhe-----	802
Db	1916	TGACCGCACTACTACCATTCGCTTCACACGCAAGCGGGCTATTGGGAGCGGCTACGAT	1975
Oy	803	-----AsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIleArg	821
Db	1976	CCACCGAGCGGCTGCCGCTTACGAGAACTACACCGCGGCTTCAACTCGGTGGGG	2035
Oy	821	LysTyrAlaHisAspSerLeuSerProIleSer-----	831
Db	2036	GCTTCAAGGACAGCAGCTGGGGCTGCTAGTACCCCCCTTAGGGCTTACAAGATGGCA	2095
Oy	832	-----AspLysGlyTyrLeuThr-----	837
Db	2096	AGATCATTTCCCTGGGCGCTGATGATGAGGTGGCTGACTACCAACCGGACCAAGATCCGAAG	2155
Oy	838	-----GlyGlyGlnValLeuAlaValGlyIthrAlaGlu-----TyrAsnTyrGlnPheM	854
Db	2156	CCCTTGGTGTAAACATCCTGATCTACCGCGCGCGCGAGCTGCTTCGCCGCTCCGCTTCG	2215
Oy	854	eLysAsp-----LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp-	870
Db	2216	TCAAGGACACCGCCAGTGGCTACCGTGCCTTCTGGAGACGTGGTAAGCACTTCGATA	2275

```

Oy      871 -----LysGlyPheThrAsnAspThrLysIle----- 879
Db      2276 CCGACTGCCCGCCACCAACACACCACTGCAGCGGCTCAAGACCGACACTTCGCCCA 2335
Oy      880 ---GlyAlaGlyAlaGlyAlaGlyTrpAlaSerProValGlyIleValAlaGlyAlaAspVala 899
Db      2336 GTTCGGTGGGCGTGGCGCTGACCTGATCACCAGCGCGTGGCCCGCTGACCTTCAGCTGG 2395
Oy      899 IaThrGlyValLysGluGluGlyAsn 907
Db      2396 CGACGGCGATCAAGAACCGCGACAC 2421

RESULT 23
US-10-419-128-14662
; Sequence 14662, Application US/10419128
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/10/419,128
; PRIOR FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: US/09/252,991
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14662
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-10-419-128-14662

Alignment Scores:
Pred. No.:      4.95e-05      Length:      516
Score:          174.50      Matches:      49
Percent Similarity: 46.28%      Conservative: 38
Best Local Similarity: 26.06%      Mismatches: 62
Query Match:      3.69%      Indels:      39
DB:              8      Gaps:      6

US-09-914-168-2 (1-919) x US-10-419-128-14662 (1-516)
Oy      230 SerIleProAlaGluAspGlnThrAlaLeuVal-----AlaAlaArgAlaVal 245
Db      23 GCGTTGCACACCTTCCGTCGCGCAATGCCAGGCGCGACGGCGAAAGCGCGGCCACGCGCTC 82
Oy      246 GLyTyTrpAspIleAspLeuSerIleLeuArgAsnSerIleGlyIleValAspValIle 265
Db      83 GCGTACTTCCAG-----GCCACAGTCAACAGCGAG 112
Oy      266 IleHisAsp-----LeuGlyGluProValTyTr 274
Db      113 GTGAAGACGGGACCGCCGCGCAAGCTCAAGCTCAAGTAGTCCCGCGCAGCCGCTTCAGCGC 172
Oy      275 IleAspTrpArgAlaValGluValArgGlyGluGlyAlaAspAspLysAlaPheThrThr 294
Db      173 CTGCCCCAGGTCAATCATCCAGGTGCTCGGCGAGGCCCGCCAGGCTGGAGAGCTTCCGCTTG 232
Oy      295 ValAlaAspGluValProLeuLeuIleGlyAspValPheHisHisGlyIleTyTrpGluThr 314
Db      233 CCGTGGCGCAAG---CAACTGAAGCGGGCGCGCAAGCTCAACCAAGGTTGTTCAAGAGGAC 289
Oy      315 LysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTrpPheAspLysAlaTrpLeu 334
Db      290 GCCAAGCGGCTGATCCAGAACCCAGGCGCTCGCCTTCTTCCAGGCGCGCTTCAGC 349
Oy      335 AspArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyTrpAsp 354
Db      350 ACCCAAGCGGCTGAGCATCCGCGCGCGCGCATTCGCCGACATCGACCTGGTCTAGAC 409

```



```

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32708
; LENGTH: 2163
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-10-282-122A-32708

Alignment Scores:
Pred. No.: 0.00442 Length: 2163
Score: 160.50 Matches: 159
Percent Similarity: 33.88% Conservative: 110
Best Local Similarity: 20.03% Mismatches: 300
Query Match: 3.40% Indels: 225
DB: 9 Gaps: 38

US-09-914-168-2 (1-919) x US-10-282-122A-32708 (1-2163)
QY 187 PheYrGlnSerSerGlnSerGlyGlnThrSerAlaIleGlySerSerHisGlnLysThr 206
Db 178 TTTAATGGCTCATCAACAGATATTAAGCTCAGCAATTTACAAAATATTGCTGTACT 237
QY 207 GluProTyrAlaAsnIleLysAlaIleuGlnAspIleThrGlnGluSerAlaMetasp 226
Db 238 GGTAAATGAGCAAAATGCAATTAATAATCAGGAGCTACTTTAGTAGTAATTAAT 297
QY 227 LeuAsnGlySerIleProArgLeuArgGlnThrAlaLeuValAla 241
Db 298 TTTAATGGCGCTAAGCTTACTAGTGAATAATGGTACAGCAATTAATGTTGAAGTAACTTT 357
QY 242 --AlaArgAlaValAlaGlyTyrTyrAspIleAspIleSerIleIleArgAsnSerIleGly 260
Db 358 CCTAAGCAGAGCAACAGGCTTATATCCGAGATGGCGCTATTATGAGCAAAAACCCGA 417
QY 261 GluValAspValIleIleHisAspLeuGlnProValTyrIleAspTyrArgAlaVal 280
Db 418 GCGATTGATTTTATGACCTTTAAAAACAGC-----TTTAGAGT--- 456
QY 281 GluValArgGlnGlyAlaAlaAspAspLysAlaPheThrThrValAlaAspLysAlaPro 300
Db 457 GATGTTAATGGGAG----- 471
QY 301 LeuLeuIleGlyAspValAlaPheHisIleGlyTyrGlnThrLysAsnLeuIleGlu 320
Db 472 ---ATCAACGGTGATATTTTAGGGAATATCATGTGGATACCAAAATCATTTTCTTACT 528
QY 321 AsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgTyrPheAspArgSerValAspVal 340
Db 529 CAGGAGGAGCACAATAATGATATTATGACGCTAT-----AAATTCACGCTGT 579
QY 341 IleLeuProAspAsnThrAla-----AspValSerLeuLeuIleYr 353
Db 580 CCTTAATTTGAAACATGCGCAATTTAACATCAAGGTAAGATAAACCATTGAATGG 639
QY 354 AspThrGlyThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspProLysThr 373
Db 640 GAAGGTAACTTCATGAATAAAGGAATAAGCCAGCTTATTTCCGCTGATGATATAATACC 699
QY 374 AsnGlnLeuThrThrAspProAspLysLeuProValLysArgGlnLeuGlnLeu 393
Db 700 AAT-----ACTGAGCAAACTATCTTACATGTACCGCGGAT----- 735
QY 394 LeuThrValAsnMetGlyGluAlaIleYrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnsp 413
Db 736 GTCCACCTTTAAAAAGTAGCCAAACCAAACTT-----CAT 771
QY 414 LeuIleAlaThrArgTyrPheAsnMetValAsnThrGlnIleValAlaPheProGlnArgGlu 433
Db 772 TTTAAGGACACAGTGTGATTAACATCATTAATAAGATATTGTATTATGCTTCAACAC 831
QY 434 GlnIleGlnAsnAspGln----- 439

```

```

; Sequence 32708, Application US/10282122A
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Treawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16

```


Db 305 GAAGATTGACGAGCAGCAACCTATTTA-----GATCCCAATCCATTAAACCA 358
 Qy 401 AlATyrAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAlaThrArgTyrPhe 420
 Db 359 ATGTCAATATATATAATATACATATACATATCCATACCTTGATA----- 403
 Qy 421 AsnMetValAsnThrGluIle--ValPheProGluArgGluGlnIleGlnAsnAspGln 439
 Db 404 -----GTGAACCTGCATATCTCCCTCTTCTGTAGAGAAATTCATC----- 445
 Qy 440 ValSerPheGluGlnSerSerSerSerArgThrGluProAlaGlnValAspLeuSerThr 459
 Db 446 -----TCT 448
 Qy 460 LeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAspLeuSer--- 478
 Db 449 CTAAAGCTAATCCCTG-----CTCTTCTGTATATATCCCTAT 484
 Qy 479 ProIleGluPheSerAlaSerAsn-----LeuIleGlnAsp----- 490
 Db 485 CCTCTGTAGTAGCTCATATGCAATATATATACATATATATGATATCCATTAGCCAAATTA 544
 Qy 491 ---LysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMetProAspArg 509
 Db 545 AAAAACTCAATAC-ATATATTAATCGATGATGATATACATTTGATATAT---GATTTA 597
 Qy 510 ValLeu-----AlaIleAsnHisAspAspGly 518
 Db 598 GTAATTAACACAGACCGCTTACTCATATGCGTAAAGTACACACCTCCACATATCTCAT 657
 Qy 519 ValAsnArgSerIleLeuGluYArgIleSerAspAlaValSerAlaValAlaArgAlaIle 538
 Db 658 ATTTCACAAATATCTCTTAATTAATTT-----GCTACATGAGAAAGAAAGAGGTA--- 708
 Qy 539 LeuProAspGluSerGluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsn 558
 Db 709 -----TGGACACATGCTACCTCTTCTTCTTA-TCGAACACGCGCTTCAACAAAC 755
 Qy 559 ArgLysThrProAlaAspValTyr-----GlnSerLysLysValPro--- 572
 Db 756 AGG-----CCTTTTCAGTCTATTAACCTCTCAATCCAGAGATCAAAAAGAACAAATGCG 809
 Qy 573 -----LeuTyr-ValPheValAlaSerAsp---LysProArgAspGlyGlnIleGlyLe 589
 Db 810 TAACACCCCTATAGAACTTTGTGCAACAGAACGCCCTCCAGAGCTATTATGCTGGCT 869
 Qy 589 uGlyTyrGlySerAspThrGlyThrArgLeuValThrLysPheGluHisAsnLeuIleAs 609
 Db 870 ACAATACTCTCTGATCAAGATGATTGCTGCACGTGGAGCTTGGAACAT----- 918
 Qy 609 nArGAspGlyTyrGlnAlaGlyAlaGluLeuArgLeuSerGluAspLysGlyValLys 629
 Db 919 -CGAATGCTTTTGTGAATGAGAACTTTTGTGATA----- 954
 Qy 629 sLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGln 649
 Db 955 -----ACAGCACCAATAGT-----CGAGATGATCAAAAATATTATGCAAA 995
 Qy 649 yTyrGlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrIle 669
 Db 966 CTTCCAAAACCAACCGCTTGCGCGTCCCAATATCATATTA---ATTAGGAACACACACT 1052
 Qy 669 uGluHisGluIleSerArgSerIleIleGlnAsn-----GlyGlyTyrPhe 684
 Db 1053 TAAAAAGAAATATACAAAAGTTACAAACACACACTGCATATTGCTTACGAATTGA 1112
 Qy 684 nArgThrTyrSerLeuArgTyrArgLeuAspLysLeuLysThrGlnAlaProGluLuth 704
 Db 1113 ACGACAAATTAATAGACGT----- 1131
 Qy 704 rTyr-----GlnAspLeuProValAsp-----PheValAsnGlyLysProSerGln 719
 Db 1132 -TGGTTTGGTACTAGCAGCTCTTTCAGTTGATACAGATTATGAGATCGAGATTCTAT 1190

Qy 719 nGluAlaLeu-----LeuAlaGlyVal-----AlaValHisLysThrValAlaAspAs 735
 Db 1191 AAAAAAATATTACTCTTTTGGCATCCCTTATCATATACAGGAGATGTTCAAGA 1250
 Qy 735 nLeuValAsnProMetArgGly----- 742
 Db 1251 TCCCTTATCTTATCTTATCCAAAGAACAAAGCTACCTTAATATGTTACTCTTATATGGTAA 1310
 Qy 743 -TyrArgGlnArg-----TyrSerLeuGluVal 751
 Db 1311 ATATAAAAAAGATTTTGCCTTACCTTACGTAGTCGGTTGATTTTACCTTTATCATAGAGCT 1370
 Qy 751 lGlySerSerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGln 771
 Db 1371 TCTTAACACAGGAACCTATCTTGCTACAAA----- 1404
 Qy 771 yValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGln 791
 Db 1404 ----- 1404
 Qy 791 yIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhe 811
 Db 1405 -ATACCAATAGGTTCCTCTCTAGGAAAGATATAGAAAACCTATCCCTCAATACTAGGTT 1463
 Qy 811 ePheAlaGlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSe 831
 Db 1464 TTATCTGGGGGTGGTGGTGGTGAAGAGGTATGACTATCATATCATTTGGACCAAAAA 1523
 Qy 831 rAspLysGlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyr-AsnT 851
 Db 1524 TAAATATGGCGATGCTATGAGAGA-----CTTCTTTTCAACAGTATAGTTGAAT 1577
 Qy 851 yTrGlu-----PheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaT 869
 Db 1578 ACGATTAAAAATTAACAGAAATCCATGCGATGTCCTCAATTTATGATGCGAGGCAATATT 1637
 Qy 869 yrAspLysGlyPheThr-----AsnAspThrLysIleGlyAlaGlyValGly 885
 Db 1638 ACGAAAAAATAATTTCTGCTACTTAAAAAATCANTA-TATTGGGGGAGTACGCGGGGC 1696
 Qy 885 aLArgTyrAlaSerProValGlyGlnValArgValAspValAlaThrGlyValLysGluG 905
 Db 1697 TACGATATATATCAAGTTTGGCCCATACGTTTGAATATACCACTCCACTTCAAGATA 1756
 Qy 905 uGlyLysn-----ProIleLysLeuHisPhePheIleGlyThrProPhe 919
 Db 1757 GAAGCCATATATAACACTTTCACCTTTATATTAGTATGCGCAAGCATTTC 1806

RESULT 29
 US-10-369-42248/c
 : Sequence 42248, Application US/10369493
 : GENERAL INFORMATION:
 : APPLICANT: Cao, Yongwei
 : APPLICANT: Hinkle, Gregory J.
 : APPLICANT: Slater, Steven C.
 : APPLICANT: Goldman, Barry S.
 : APPLICANT: Chen, Xianfeng
 : TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
 : FILE REFERENCE: 38-10(52052)B
 : CURRENT APPLICATION NUMBER: US/10/369,493
 : PRIOR FILING DATE: 2003-02-28
 : PRIOR APPLICATION NUMBER: US 60/360,039
 : NUMBER OF SEQ ID NOS: 47374
 : SEQ ID NO 42248
 : LENGTH: 1890
 : TYPE: DNA
 : ORGANISM: Halobacterium sp. NRC-1
 : US-10-369-493-42248

Alignment Scores:

```

Pred. No.: 0.00718 Length: 1890
Score: 157.00 Matches: 99
Percent Similarity: 40.00% Conservative: 51
Best Local Similarity: 26.40% Mismatches: 133
Query Match: 3.32% Indels: 92
DB: 8 Gaps: 20

US-09-914-168-2 (1-919) x US-10-369-493-42248 (1-1890)
Oy 248 TTTGACCTTCATGCCAGCTCAGCGGCTCAGCTCACCACACGATGCTC----- 1078
Db 1131 TTTGACCTTCATGCCAGCTCAGCGGCTCAGCTCACCACACGATGCTC----- 1078
Oy 268 AATPLeu---GluGluProValTyrIleAspTyrArgAlaValGluValArgGlyGluGly 286
Db 1077 GACCTCCCGGAGACGACGCCACCTGAGCGCGCGCGGAGCGGACCGCCCTGCTCGG 1018
Oy 287 AAlaAspLysAlaPheThrThrValAlaAspGluValProLeuLeuIleGlyAspVal 306
Db 1017 GTT---GACGTTCTTCTGGGCTGCTGCGGAGAGAGCTCTGACCTGCGCGCTGAGCTG 961
Oy 307 PheHisIleGlyLysTyrGluThrLysAsnLeuIleGluAsnAlaSerAlaGluHis 326
Db 960 TGGCATTCTCGTGACCCCGGACGAGATCAGCTCGATGCTCTTTGCGAGAGCC 901
Oy 327 GLYTyr-PheAspGly-----ArgTrpLeuAspArgSerValAspValIleLeuProAsp 344
Db 900 GCGCTCTCGAGGCGCTGCTGCGCGCGCGGAGCGGAGCTGATGAGTCTCGGCTGAT 841
Oy 345 -----AsnThrAlaAspValSerLeu---IleTyrAspThrGlyThrGlyTyr 359
Db 840 GCTCTCGAGAGCTGACGCGCTGATGCTGCTGAGAGTACCGCGCGGAGTGGCTGCG 781
Oy 360 ArgPheAspGluValAlaPhePheThrIleAspProLysThrAsnGluLeuThrTrpAsp 379
Db 780 CGTGACGACGACGAGTT-----GAC 760
Oy 380 ProAspLysLeuProValLys-----ArgGluLeu 389
Db 759 CGTCTGCTCTCGGCTCGACAGCTGATTTGGCTTCTCGCGCGGCTCGCTGAGTGC 700
Oy 390 LeuGluGluLeuLeuThrVal---AsnMetGlyGluValArgLysLeuGluAlaValArg 408
Db 699 CTGAGCGCGCTGCGGCTCGCGGAGTGCATCCCTGCTCTCGCAAGTTGCTCGC 640
Oy 409 AlaLeuSerAsnAspLeuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleVal 428
Db 639 GAGGTGCTGATGATGCGCGTCCCA-----GTCGCGCGGACGAGTCTGTT 592
Oy 429 PheProGluArg-----GluGlnIleGluAsnAspGluVal 440
Db 591 GTCCCTCTGCTGCTGCGACCTGTAAGACCGCGCGCGGAGTCCGAGTATCGACAGTGC 532
Oy 441 SerPheLeuIleSerSerSerArgTrpGluProIleGluValAlaIleAspThrLeu 460
Db 531 -----GAACTGCGCGCCACCGAGCTGTAAGCCAAACAGCTG 493
Oy 461 GluProVal---IleGluThrValGluLeuThrAspGlyIleLeuMetAspIleSerPro 479
Db 492 GTCCGCTCTGCTGCGAGCGCGTATGCGCATGACGCGCGC----- 454
Oy 480 IleGluPheSerAlaSerAsnLeuIleGluAspLysLeuAsnLeuValAlaIleAlaVal 499
Db 453 -----CCTCGCTGCTGAGCATGCTGCTGACAGCTC-----GAAAGCC 418
Oy 500 ArgHisLeuTyrAspMetProAspAspArgValLeuAlaIle-----Asn 514
Db 417 CGCGATCTGCGCGGCTGCTGCTGCTGCGCGCTGCGCTGTAAGTACGCGCGGAC 358
Oy 515 HisAspAspGlyValAlaAsnArgSerIleLeuGlyArgTyrIleSerAlaValSerAlaVal 534
Db 357 CGTGATACGCGC-----TTTCTGACGCTCTGCGCGGAGATACCTCTCGCGCTC 310

```

```

Oy 535 AlaArgAlaIleLeuProAspLysSerGluAsnGluValIleAspLeuProGluArgThr 554
Db 309 GCGTTGATGCTCTGGAGATGCGGCGCA-----GATCTCTCGCGGCTGA 262
Oy 555 AlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValProLeuTyr 574
Db 261 CTTGCTC-----GTCGCGGAGGCAACCGTGTACTGCTCTC 226
Oy 575 ValPheValAlaSerAspLysProArgAspGlyGlnIleGlyLeu 589
Db 225 GCCCATGTGCGCTGATGCA---CCGATGTGCTGATGCGGATT 184

RESULT 30
US-60-453-134-135
; Sequence 135, Application us/60453134
; GENERAL INFORMATION:
; APPLICANT: BAKALETZ et al.
; TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF H. INFLUENZAE
; FILE REFERENCE: 28335/3815
; CURRENT APPLICATION NUMBER: US/60/453,134
; CURRENT FILING DATE: 2003-03-06
; NUMBER OF SEQ ID NOS: 588
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO: 135
; LENGTH: 2349
; TYPE: DNA
; ORGANISM: H. Influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (553)..(553)
; OTHER INFORMATION: n = a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (2238)..(2238)
; OTHER INFORMATION: n = a, c, g, or t
US-60-453-134-135

Alignment Scores:
Pred. No.: 0.0167 Length: 2349
Score: 154.00 Matches: 135
Percent Similarity: 31.37% Conservative: 94
Best Local Similarity: 18.49% Mismatches: 239
Query Match: 3.26% Indels: 262
DB: 10 Gaps: 32

US-09-914-168-2 (1-919) x US-60-453-134-135 (1-2349)
Oy 268 AspLeuGlyGluProValTyrIleAspTyrArgAlaValGluValArgGlyGlu----- 285
Db 119 GATGTAATGATAGGTTTACGATGATGATGATGATGATGATGATGATGATGATGATG 178
Oy 286 GlyAlaAspLysAlaPheThrThrValAlaAspGluValProLeuLeu----- 302
Db 179 GGTATGCT-----GCCGAGCTTGAACTTACTTACTTACGACATTA 217
Oy 303 ---IleGlyAspValPheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGluAsn 321
Db 218 CATTTAATGATCTCTCCCGCTAGTATGATGATGATGATGATGATGATGATGATGATG 277
Oy 322 AlaSerAlaGluHisGlyTyrPheAspGlyArgTrpLeuAspArgSerValAspValIle 341
Db 278 AAATCTGGAACCGGCTTAC-----GTTAGCGCAACGGTAAATTCAGTA----- 322
Oy 342 LeuProAsp-----AsnThrAlaAspValSerLeuIleTyrAspThrGly 356
Db 323 ---CCTGATTTGATGATGCAAAATTAACATTAACATTAACATTAACATTAACATTA 379
Oy 357 ThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspProLysThrAsnGluLeu 376
Db 380 CGAGGTTTAAGTGTGCGCAACTTCTGCTT-----GAGGAAATACCGCTT 424
Oy 377 ThrTrpAspProAspLysLeuProValLysArgGluLeuLeuGluGlnLeuLeuThrVal 396

```

Db 425 TCTGCTGAT-----AGCACTTTACGTCAGAAATGCCCAA 460
 QY 397 AsnMetGlyGluAlaIleuValAsnLeuGlnAlaValArgAlaLeuSerAsnAspLeuIleAla 416
 Db 461 CAAGAAAGAACTTGATATATTCACATTAAGT----- 493
 QY 417 ThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgGluGlnIleGln 436
 Db 494 -----GAGTTAGAGAAATTCGC 511
 QY 437 AsnAspGlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAsp 456
 Db 512 TTAGATCGTACAGCTTCTTCGAAACAGCGAAACCGAATNGATCTCT----- 559
 QY 457 GluSerThrLeuGluProValIleGluThrValGluLeuThrAspGlyIleLeuMetAsp 476
 Db 559 ----- 539
 QY 477 IleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspPlysLeuAsnLeuValAla 496
 Db 560 -----ATCAATGCTAGCAAT-----GATGAAGTGATGCTGA 592
 QY 497 AlaLysAlaArgHisLeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAsp 516
 Db 593 -----TATAAAGTCMAAGACGTAACACGGGATCACTTCTT----- 631
 QY 517 AspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArg 536
 Db 632 ---GGATTTGCTTACCGTACAGAGAGTGGTATTAGTTATACAGCACTGTT----- 679
 QY 537 AlaIleLeuProAspGluSerGluAsnGluValIleAsnLeuProGluArgThrAlaLeu 556
 Db 680 -----AAACAAGATAATTTCTTAGAACAGCGCTCGAGTAAGTATA 721
 QY 557 AlaAsnArgLysThrProAlaAspValTyrGlnSerLysValProLeuTyrValPhe 576
 Db 722 GCTGGTACGAAA----- 733
 QY 577 ValAlaSerAspLysProArgAspGlyGlnIleGlyLeuGlyTyrGlySerAspThrGly 596
 Db 734 -----AATGATTATGCT 745
 QY 597 ThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGly 616
 Db 746 ACCAGCTCATTTGGGTATACCGAGCCCTATTCTTAAGATGCTGAAGCTTGGT 805
 QY 617 AlaGluLeu-----ArgLeuSerGluAspLys 626
 Db 806 GGAATGTTTTCTTTGAAACATACGATACCTTAAGATGATACATCCCTACACTATAG 865
 QY 627 GlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAla 646
 Db 866 CGTACGACTATATGGAAGTATGTTACTTATAGCTTCCCTGAATGAATACCTCTAT 925
 QY 647 ThrLeuGlyTyrGlnGlnGlnValPheGlyHisSerThrAsnGlyPheAspLeuSerThr 666
 Db 926 TATGTAGGA-----TTAGGCCATACCTATATAT-----AAATATAGTAAC 964
 QY 667 ArgThrLeuGlnHisGluLeuSerArgSerIleIleGlnAsnGlyGlyTyrPAsnArg 685
 Db 965 TTTCCTCTAGAAATAT-----ACCCGTAATTTATATATTTCAA----- 1000
 QY 686 ThrTyrSerLeuArgTyrArgLeuAspLysLeuThrGlnAlaProProGluThrTyr 705
 Db 1001 -----TCAATGAATTTAAAGTAATGCAATTAACAAAT----- 1036
 QY 706 GlnAspLeuProValAspPheValAsnGlyLysProSerGlnGlnAlaLeuAlaGly 725
 Db 1037 -----GACTTGTATTTTCTTTTGGTTGG----- 1060
 QY 726 ValAlaValHisLysThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGln 745
 Db 1061 -----AACTATAACAGCCTTATATAGAGCTATTTCCTCA 1093

QY 746 ArgTyrSerLeuGluValGlySerSerGlyLeuVal-----SerAspAlaAsn 761
 Db 1094 ACTAAAGGGGTAAAGCAAGTCTTGGTGGAGGATTACTTTCAGGTTCTGTATACAA 1153
 QY 762 MetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGly 781
 Db 1154 TACTACAAACTAAGTCAGATGTACAGAGGTTTCTACCCATTA----- 1195
 QY 782 SerAsnArgAlaHis-----GlnMetThrGlyGlyIleGlnAlaGlyTyrIleTyrSer 799
 Db 1196 ---GACGAGATCAACCTCTGGGTGTATCTGCAAAACATCTCAGAGTATGCAAAATGCT 1252
 QY 800 AspAsnPheAsnHisValProTyrArgLeuArgPhePheAlaGlyGlyAspGlnSerIle 819
 Db 1253 TTGGAAACAAGCGTTTACCGCTTCTATCAACACTTATACACGGGTGCACTGTTCA 1312
 QY 820 ArgGlyTyrAlaHisAspSerLeuSerPro----- 839
 Db 1313 CGTGGTTTTCCTTATGTAGTACTATGTGACCTAACGCAATTTATGCCCTCAAAATGCTAAT 1372
 QY 830 -----IleSerAspLysGlyTyrLeuThrGly 838
 Db 1373 AGTAATAGTAATAGTAATGTACTTTTAATAGATAGTCTGAT-----GTGATTGCT 1426
 QY 839 GlyGlnValLeuAlaValAlaGlyThrAlaGlu-----TyrAsnTyrGluPheMetLysAsp 856
 Db 1427 GGTATGCAATCAACACTCGGAGTGGCAGAACTTATTTACCACACCGCTTTGGAGCAT 1486
 QY 857 -----LeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyr----- 869
 Db 1487 AAGGCCAAATACGGTCCGAACTCCCTATTGTTGATGCGCAAGGTGTAATACT 1546
 QY 870 -----AspLys 871
 Db 1547 AAGTGAATCGATTAAGACTGGGTAGATTAACAGTATTAACCTTAACCGCATTAAT 1606
 QY 872 GlyPheThrAsnAspThrLysIleGlyValGlyValAlaArgTyrPheAspProVal 891
 Db 1607 GCGCAATCAAGCCGTATTCGCCGCTCTACAGGTGTGCGATTCCCAATGCAATCTCCATAT 1666
 QY 892 GlyGlnValArgValAspValAlaThrGlyValLys-----GluGluGlyAsnProIle 909
 Db 1667 GGGCCATTTGCTATCTCTTATGCCAAACCAATTAATAATATGATGATGTGCA 1726
 QY 910 LysLeuHisPhePheIleGlyThrProPhe 919
 Db 1727 CAGTTCCATTTAGTATTGAGGTTCTTTC 1756

RESULT 31
 US-10-282-122A-27975
 : Sequence 27975, Application US/10282122A
 : GENERAL INFORMATION:
 : APPLICANT: Wang, Liangsu
 : APPLICANT: Zamudio, Carlos
 : APPLICANT: Malone, Cheryl
 : APPLICANT: Heselbeck, Robert
 : APPLICANT: Ohlsen, Karl
 : APPLICANT: Zyskind, Judith
 : APPLICANT: Wall, Daniel
 : APPLICANT: Trawick, John
 : APPLICANT: Carr, Grant
 : APPLICANT: Yamamoto, Robert
 : APPLICANT: Forsyth, R.
 : APPLICANT: Xu, H.
 : TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 : FILE REFERENCE: ELIPIRA.034A
 : CURRENT APPLICATION NUMBER: US/10/282,122A
 : CURRENT FILING DATE: 2003-02-20
 : PRIOR APPLICATION NUMBER: 60/191,078
 : PRIOR FILING DATE: 2000-03-21
 : PRIOR APPLICATION NUMBER: 60/206,848
 : PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patent version 3.1
SEQ ID NO: 27975
LENGTH: 3111
TYPE: DNA
ORGANISM: Mycoplasma pneumoniae
US-10-282-122A-27975

Alignment Scores:

Pred. No.: 0.0241 Length: 3111
Score: 154.00 Matches: 141
Percent Similarity: 37.05% Conservative: 98
Best Local Similarity: 21.86% Mismatches: 248
Query Match: 9 Gaps: 159
DB: 29

US-09-914-168-2 (1-919) x US-10-282-122A-27975 (1-3111)

QY 3 LysProValLeu-----PheAlaAsnArgSerPhe-----MetProValAla 16
DB 1138 GAGCAATTATTAAAGTTGTTGTTAAATGAAGATTGTTCAAAAACCAACCCAA 1197
QY 17 LeuAlaIleArgLeuProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAsnProAla 36
DB 1198 ACAGCAGCGCAGCTAGTCAAGCCCTCTCAAGCAATTACATTGCTTTAAATGAAGCT 1257
QY 37 AsnIleIleAsnHisValProAlaHisAspThrAlaIleAsnGlnAlaValAlaGlyAsn 56
DB 1258 GACCTCTTCATCATTAATGAGGCC----- 1281
QY 57 ProProValLeuLeuThrProGluGlnIleGlnAlaArgLeuAsnAlaIleGlyLeuAsn 76
DB 1282 -----GTTGGCAGCCAGCCAGCCAGCA 1302
QY 77 AlaLysProGlnSerGlnAlaLeu---AspValValAsnPheAspAspGlnSerProIle 95
DB 1303 GCCTTACACAAAGCGAGCTCATTCAGAGAGTGTGATGACAGCGCAAGCC--- 1359
QY 96 SerArgIleGlyGlnSerProProLeuGlyLeuAspMetSerValIleGlyGlnThr 115
DB 1360 -----CAACAGAGCTGGGTTTCATTATGATGATGATTTCTCAGTACT 1404
QY 116 ThrProLeuSerLeuGlnGluLeuPheAlaGlnGlnSerThrGlnMetGlyIleAsnPro 135
DB 1405 GCACCACTACTGTAGCGAAGATAGAGTTCAGACAGCAAGAGCTAGTC----- 1452
QY 136 AsnAspTyrIleProGluTyrGlnGlyGlu-----GlnProAsnSer 149
DB 1453 AATGATTTGTACAACTACTAGTAGAAGAAACAACACTACTTTGCATCAACTCCAGTTT 1512
QY 150 GluValValValProProThrLeuGluProGluLysProGlyLeuIleLysArgLeuTyr 169
DB 1513 GAACACCTTTGTAATTCACAGTAGAAGTAATCTGACAGAACACTTTTGAAGAACAGAGTTT 1572
QY 170 AlaArgLeuPheAsnAspGlyValAsnLysValProArgLeuLysAlaLysPheTyrGln 189

DB 1573 GAATCACTGTTGTAAGTGCACATCAATGAACCTAATGTT-----GCA 1617
QY 190 SerSerGlnSerGlyGluThrSerAlaIleGlySerSerHisGlnLysThrGluProTyr 209
DB 1618 TCACTCCAGTTGTTGAACGTTGAACCTTACAGAAACACCCAGTAAGTTTGAAGCC--- 1674
QY 210 AlaAsnIleLysAlaIleGlnGluAspIleThrGlnGlnSerAlaMetAspLeuAsnGly 229
DB 1675 -----CTTGAACCTGTTCACTAGAAACAGCA----- 1701
QY 230 SerIleProArgLeuArgGlnThrAlaLeuValAlaAlaGlnAlaValGlyTyrTyrAsp 249
DB 1702 -----CCAGTTGTTACGAACTGTTACACTGAGAAAGAGTA-----GACCA 1749
QY 250 IleAspLeuSerIleIleArgAspSerIleGlyGluValAspValIleIleHis----- 267
DB 1750 GAAGTATTAGCAGTTGTTGAAGAGACGACCTGCTGCTAGAACCAATCTTGAACCTCA 1809
QY 268 -----AspLeuGlyGluProValTyrIleAspTyrArgAlaValGluVal 282
DB 1810 ACACATTAAGCTCCAGAGACCGTTGAGAGCGCAAGTTGACAAAGAGTCAACAGCTGTT 1869
QY 283 ArgGlyGluGlyAla-----AspAspLysAlaPheThrThrValAlaAspGluVal 299
DB 1870 GCAGTCGAACCGATATTGTAAGAACCGATCAAGCACTTCTGGAAGCCCAAGCTGACTG 1929
QY 300 -----ProLeuLeuIleGlyAspValPheHisGlyLysTyrGluThrLysLysAsn 317
DB 1930 GATTGGAGGCGTTATTTGCTAAC---AGTGAATAGCGTTACTTGAAGCAAGCAAAAC 1986
QY 318 LeuIleGluAsnAlaSerIleAluHisGlyTyrPhe-----AspGlyArgTyrLeuAsp 335
DB 1987 TCGATTGTAAC-----GTTACTTCGATGAAGATTAACAATGGGTA--- 2028
QY 336 ArgSerValAspValIleLeuProAspAsnThrAlaAspValSerLeuIleTyrAspThr 355
DB 2029 -----TCCACAGCTACAGCCCA----- 2046
QY 356 GlyThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspProLysThrAsnGln 375
DB 2047 ---ACAGCAAAATCTGAAGAGTTGTT---TTACAGCAGAGCCGCTGAACCTCAGAG 2100
QY 376 LeuThrThr-----AspProAspLysLeuProValLys-ArgGluLeuGlnGluLe 393
DB 2101 TTAAATACGTAATCTGATCTTACTTGAACCGAGTAGAATTAACCTGACCAAGCA 2160
QY 393 uLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValArgAlaLeuSerAsnAs 413
DB 2161 CTTAAGCTTTGATTGGAAGAGATTCCTGAACCGAATCAATTAATTAACCTGACCAAG 2220
QY 413 PheuIleAlaThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluArgG 433
DB 2221 CTTA-----ATTTTGA-CCAGAAAGTAGA 2243
QY 433 uGlnIleGlnAsnAspGlnValSerPheGluGlnSerSerSerArgThrGluProAl 453
DB 2244 GCTTGAACCAAGAAATTAATGAACTTCAATCGAATCGAAGAGTCAACAGAGCTAGC 2303
QY 453 aglnValAsp-----GluSerThrLeuGluProValIleGluThrValGluLeuThrAs 471
DB 2304 ACAAGAAATCAAGTTTCAGAGCTGACAGCAAGAACCTTAATTTGAACA----- 2349
QY 471 pGlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAsp 491
DB 2350 -----GAACTAGAAAGTTCAACCTGAATCAAGAAATTTGAATCTAA 2387
QY 491 sLeuAsnLeuValAlaAlaLysAlaAlaGlnHisLeuTyrAspMetProAspArgVal 511
DB 2388 ATTCGAAGCC-----GAAGTTCAATTCAGAGCCAAAGT 2420
QY 511 uAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIleSerAspAlaVal 531
DB 2421 AAGTTTGAATTCAGATTTTGAACCTAA-----CCAGAAAGCTCA 2459

Oy	484	AlaserAsnLeuLleGlnAspLysLeuInsValAla	-----AlaLysAla	499
Db	1513	AAATTCGACGAATTGGAAAAACGTTTGGCAGACGGCAGCAACGCGAGAAAGC	1513	499
Oy	500	ArgHis	-----	501
Db	1573	CATCATATGAACCTGCTGCTAACAAAGTAACGTATGCGCAAAATTCCTTGGT	163	
Oy	502	-----LeuTyrAspMetProAspArgValLeuAla	512	
Db	1633	CGCTGGACGGGATTCCTGTATCCAGAAATGTTGGAGACTGAAAAAGATAACTTCTAGC	168	
Oy	513	ILeasnHisAspAspGlyValAsnArgSerLleLeuGlyArgLleSerAspAlaValSer	533	
Db	1693	ATGGCGCATGAA-----TTACATAAACCTGTCACTCGGGCAA---AATGAGCGTGTGAG	174	
Oy	533	AlaValAlaIArgAlaLleLeuProAspLysSerLysnGluValILeAspLeuProGlu	555	
Db	1744	GGCTGATCTAATGCAT-----	176	
Oy	553	ArgThrAlaLeuAlaAsnArgLysThrProAlaAspValTyrGlnSerLysValPro	573	
Db	1762	CGTCCGATCCCGTGGCGGATTTATCCGATGCA-----AATCGCCGATGTGT	180	
Oy	573	LeuTyrValILePheValAlaSerAspLysProArgAspGlyGlnLleGlyLeuGlyTyrProLys	593	
Db	1807	TCTTTATGCTTTTG-----GGGCCACGCGGGGTAGT-----	183	
Oy	593	SerAspThrClyThrArgLeuValThrLysPheGlnHisAsnLeuILeAsnArgAspGly	612	
Db	1840	-----AAACCGAATGTGGCAAGCGGTGGCCAAATTCCTATTTGACAGACGTAT	188	
Oy	613	TyrGlnAlaGlyAlaGlnLeuArgLeuSerLysAspLysGlyValLysLeuTyrAla	633	
Db	1888	---GATGGCATGGTGGCCCTTGGACATGTGGAAATTTATGGA-----	192	
Oy	633	ThrLysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrGlnGln	655	
Db	1927	-----AAACATACCTGTTTCAAGACTGGTGGGGCTCCTCTGGATAT	196	
Oy	653	GluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuHisGlu	673	
Db	1969	-----GTTGCATATGAAGAGCGCGGTAT-----CTGACAGAAAGC	200	
Oy	673	IleSerArgSerIleLeuGlnAsnGlyLysTyrAsnArg	685	
Db	2005	GTGGCTGCTGCCCTTATTCAGTATTTTACTGGATGAGATAGAAAAAGCTCATCTGAT	206	
Oy	686	-----ThrTyrSerLeuArgTyr	695	
Db	2065	GTCTTCACACATTCGTTCACAGCTATTTGGATGATGCTGCTGACGAGTGGACAGC	219	
RESULT 33				
DS-10-282-122A-33378				
Sequence 33378, Application US/10282122A				
GENERAL INFORMATION:				
APPLICANT: Wang, Liangsu				
APPLICANT: Zamudio, Carlos				
APPLICANT: Malone, Cheryl				
APPLICANT: Haselbeck, Robert				
APPLICANT: Ohlsen, Karl				
APPLICANT: Zyskind, Judith				
APPLICANT: Wall, Daniel				
APPLICANT: Trawick, John				
APPLICANT: Carr, Grant				
APPLICANT: Yamamoto, Robert				
APPLICANT: Forsyth, R.				
APPLICANT: Xu, H.				
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms				
FILE REFERENCE: ELTRA 034A				
CURRENT APPLICATION NUMBER: US/10/282,122A				
CURRENT FILING DATE: 2003-02-20				
PRIORITY APPLICATION NUMBER: 60/191,078				

```

1  PRIOR FILING DATE: 2000-03-21
2  PRIOR APPLICATION NUMBER: 60/206,848
3  PRIOR FILING DATE: 2000-05-23
4  PRIOR APPLICATION NUMBER: 60/207,727
5  PRIOR FILING DATE: 2000-05-26
6  PRIOR APPLICATION NUMBER: 60/230,335
7  PRIOR FILING DATE: 2000-09-06
8  PRIOR APPLICATION NUMBER: 60/233,347
9  PRIOR FILING DATE: 2000-09-09
10 PRIOR APPLICATION NUMBER: 60/242,578
11 PRIOR FILING DATE: 2000-10-23
12 PRIOR APPLICATION NUMBER: 60/253,625
13 PRIOR FILING DATE: 2000-11-27
14 PRIOR APPLICATION NUMBER: 60/257,931
15 PRIOR FILING DATE: 2000-12-22
16 PRIOR APPLICATION NUMBER: 60/267,636
17 PRIOR FILING DATE: 2001-02-09
18 PRIOR APPLICATION NUMBER: 60/269,308
19 PRIOR FILING DATE: 2001-02-16
20 Remaining Prior Application data removed - See File Wrapper or PALM
21 NUMBER OF SEQ ID NOS: 78614
22 SOFTWARE: PatentIn version 3.1
23 SEQ ID NO 33378
24 LENGTH: 2763
25 TYPE: DNA
26 ORGANISM: Pseudomonas syringae
27 US-10-282-122A-33378

```

Alignment Scores:	
Pred. NO. :	0.0439
Score:	150.00
Percent Similarity:	32.65%
Best Local Similarity:	18.33%
Query Match:	3.17%
OB:	9
	Length: 2763
	Matches: 187
	Conservative: 146
	Mismatches: 355
	Indels: 332
	Gaps: 48

US-09-914-168-2 (1-919) x US-10-282-122A-33378 (1-2763)

[illegible]

Oy 184 LysAlaIalysPheTyrgInserSerInserGlyIuThrSerAlaIleGlySerSerHis 203
 Db 571 CCCCC-----GTTGGCGAATAAACCGCTGTGGC----- 600
 Oy 204 GlnIystrhgIuProTyraAlaAnIleLysAlaAlaLeuGluAspIleThrgInGInIuser 223
 Db 601 CTGCTGGTGGGGCGGTGGCGGATCAAGAGCTTTACGACAACTGGAAATGGTGGCG 660
 Oy 224 AlaMetAspLeuAnGlySerIleProArgLeuArgGlnThrAlaLeuValAlaAlaArg 243
 Db 661 ACCCTGGCGATCCGGGGCGC-----AAGACCTTGGCCCACTTCGAAAGACACCGC 714
 Oy 244 AlaValGlyTyraAspIleAspLeuSerIleIleArgAsnSerIleGlyIuValAsp 263
 Db 715 GCGATGGCTTCTCTCTTACGATTTGGCGACCATCAAG-----ATCGAG 759
 Oy 264 ValIleIleHisAspLeuGlyIuProValTyraIleAspTyraArgAlaValGluValArg 283
 Db 760 GTG-----CCGCTGGACATCGACCTGGACAGATTCGATTGC--- 795
 Oy 284 GlyIuGlyAlaAspAspIyAlaPheThrThrValAlaAsp---GluValProIleuLeu 302
 Db 796 AGCGAGCCAGACCGCGAGCAAGCTGATGAACCTGACGCCGCGATCGAAATTCAGAGCTGG 855
 Oy 303 IleGlyAspValPheHisIleGlyIystrhgIuThrLysLysAsnLeuIleGluAsnAla 322
 Db 856 ATCGAAGACCTCAACAGCGATGCCAAAGCGGACGCGACAGAGCTGGTGTCCAGAGCGG 915
 Oy 323 SerAla-----GlnHisGlyTyraPheAspGly 331
 Db 916 ACCGTGGCAGCCAGAAAGACAGCTGACGAGTCTTGTGACAAAGCGAGTTTCAGCGGC 975
 Oy 332 ArgTyraLeuAspArg-----Ser 337
 Db 976 ---TGCTGAATAAGCTTACGGCGCACCGCTTTGCGCTTGACCCAGACAAAGCGC 1032
 Oy 338 ValAsp-----ValIleLeuProAspAsnThr 346
 Db 1033 ACCGACGCCACAGCTGCGCACTGGTGGGCTGTCTTGGCATCCAGCCCAAGCGC 1092
 Oy 347 AlaAspValSerLeuIleTyraAsp---ThrgIyThrgInTyraArgPheAsp---GluVal 364
 Db 1093 GCGTCACTCCCGCTACCCACTCTATATATGAGCGGCCACACCTGACCTGGCGCACCC 1152
 Oy 365 ValPhePheThrIleAspProIystrhgInGlnLeuThrThrasProAspIyLeuPro 384
 Db 1153 GTCTCAAGACGCTCAAGCG-----CTACTGGAAGACCCGAAACAGATCAAG 1200
 Oy 385 ValLysArgGlu-----LeuLeuGluGlnLeuLeuThrValAsn 397
 Db 1201 GTGGGTACAGACGCCAAATTTGCCATCAACCTGCTGCCAATGGCGGATGGCGGTAT 1260
 Oy 398 MetGlyGluAlaTyraAsnLeuGlnAlaValArg-----AlaLeuSer 411
 Db 1261 CAGCGCGAGGCGATCGAGCTGCGAGGTGTGCGCTTTGACACCATTTCTGATCACTACGTA 1320
 Oy 412 AsnAspLeuIleAlaThrArgTyraPheAsn-----MetValAsn 424
 Db 1321 CTGGACTCCACCGCACTGCGACAGACGCGACAGCTGTGGGCCAATGACTGACGCAAT 1380
 Oy 425 ThrGluIleValPheProGlu-----ArgGluGlnIleGlnAsnAspGln 439
 Db 1381 ACGCGGATCACTTCACAGACATCGCGCGGACAGGGTGCACACAGCTGACCTTCGACGAG 1440
 Oy 440 ValSerPheGluGlnSerSerSerArgThrgInProAla----- 453
 Db 1441 ATCGCATCAAGCAGCGCGTAACTACGCGCGGAAGAGACCTGACCTGACCTGCGCGCTG 1500
 Oy 454 ---GlnValAspGluSer-----ThreGluIuProValIleGluThr 466
 Db 1501 CATGAAGTCTTCGAAGCCCGCTGGCGCGGATTCGACCTTCGACCGGCTGTCAACGAG 1560

Oy 467 ValGlu-----LeuThrAsp 471
 Db 1561 ATCGAGATGCCCATTTGTCGCGGTGCTGGCCCGCATCGAAAGTCAGAGCGCGCTGATCAT 1620
 Oy 472 GlyIleLeuMetAspIleSerProIleGluPheSerAlaSerAsnLeuIleGlnAspLys 491
 Db 1621 GCCAACCCTGTGGCATCTCAGAGCGTTGAG-----CTGGCGGACAG 1662
 Oy 492 LeuAsnLeuValAlaAlaLysAlaArgHisLeu-----TyraSerMetProAsp 507
 Db 1663 ATGACCGCAGCTGAGCGCGAGCGGTTCGCCATTCGCCGTGAAGAAATTCACCTCGGTTGC 1722
 Oy 508 AspArgValIleuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGlyArgIle 527
 Db 1723 CCCAAGCAGCTTGGCGGTATTTCTACGAAAACTCGGATCCGATCTCTGACGAAGAC 1782
 Oy 528 SerAspAlaValSerAlaValAlaArgAlaIleuProAspGluSerGluAsnGluVal 547
 Db 1783 GCCACCGCGCACCATTCGACAGCCGAGGGGTGTGGCGGACCTCGCGAAGACAGAGAC--- 1839
 Oy 548 IleAspLeuProGlu-----ArgThrAlaLeuAlaAsnArgIystrhgPro 562
 Db 1840 TTCCTCGTCCAGAGTGTCTGATGCAATACCTGGATGAGCAAGCTCAAGACCTAT 1899
 Oy 563 AlaAspValTyrgInserLysLysValProLeuTyraValPheValAlaSerAspLysPro 582
 Db 1900 ACCGAC-----CGCTGCCCGAGCAGATC-----AACCG 1929
 Oy 583 ArgAspGlyGlnIleGlyLeuGlyTyrgIySerAspThrGly---ThrArgLeuValThr 601
 Db 1930 CCGACCGGCGGTATTCACACCTCTATATCAAGCGCGTACCGCGGTGACCGGGCTTGTCT 1989
 Oy 602 LysPheGlnHisAsnLeuLeuAsnArgAspGlyTyrgInAlaGlyValGluLeuArgLeu 621
 Db 1990 TCCAGCAGCCGCAATCTGCAGAAC-----ATCCCTATTGTGTACC 2028
 Oy 622 SerGluAspLys-----LysGlyValLysLeuTyraAla 632
 Db 2029 GCCGAAGCGCGCGCATCGTCAGCGGTCTGTGGCGCCCAAGGTTACAACTGCTGGCG 2088
 Oy 633 Thr-----LysProIleu 636
 Db 2089 GCAGACTATTGCGAGATTGAACTGGCGGATCATGGCGCACCTGGCCAAAGCACAGCGTTGG 2148
 Oy 637 SerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyrgInGlnGlnIuValPheGly 656
 Db 2149 CTGCATGCGTTCGCAACACCTGGACCTGCACCGCGCACAGCGTGGAAATTTTTCGCG 2208
 Oy 657 HisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSer 676
 Db 2209 GTCGAAGCTGAGACAGGTGACACCGCATGCGCGGAGGCCAAAGCATTCATTCGCG 2268
 Oy 677 IleIle-----Gln 679
 Db 2269 CTGATCTACGGCATGACGCCATTTGGCGCTGGCCAAACAGATCGCGGTGACCGCAAGCAG 2328
 Oy 680 AsnGlyGlyTyraAsnArgThrTyraSerLeuArgTyra-----ArgLeuAsp 694
 Db 2329 TCGCAGCGTATGTGATCGATCGTATTGTCCTTACCCCGCGGTGTGGCATATATGGAG 2388
 Oy 695 LysLeuLysThrGlnAlaProProGluThrTrpGlnAspLeuProValAspPheValAsn 714
 Db 2389 CGCAGCGCGCACCGCGCGCGCAGCAG-----GGCTTCGTGCGAG 2427
 Oy 715 GlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAsp 734
 Db 2428 ACTATTTTGGCGGTTCGCTGATCTGACCGCGCAGATC----- 2463
 Oy 735 AsnLeuValAsnPro---MetArgGlyTyraGlnArgTyraSerLeuGluValGlySer 753
 Db 2464 AACGCCAAGACCTTCACTGCGCAAGCGCGCAAGCATGCGCATCAACCGCGCATG 2523
 Oy 754 SerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyra 773

```
Db 2524 CAGGCGACCGCGCGAC-----ATCATCAAAAAAGCATGTGGCGGTGAAT 2571
Oy 774 SerpHeGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyLeuGln 793
Db 2572 GCGTGGCGTGCAGCAATCC-----GCGCTGTGAT 2598
Oy 794 AlaGlyTyrLeuTyrPheSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAla 813
Db 2599 GCACGGGTCACTGTCACAGCTGCAC----- 2622
Oy 814 GlyGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAsp--- 832
Db 2623 -----GATGAACCTGGTACTCTGGAAGCTGCCGAGACCTGGTTGATCAGATCAGCGAGCAG 2676
Oy 833 ---LysGlyTyrLeuThrGlyGlyGlyValValLeuAlaValGlyThrAlaGlyTyrAsnTyr 851
Db 2677 ATCCGTTCACACATGAGCGGAGCGGACAGACTGGCAGTG----- 2715
Oy 852 GluPheMetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLys 871
Db 2716 -----CCGCTGCTGCTGGAAGTCCGGGTGCGGCAACACCTGGGAGCAG 2757

RESULT 34
US-10-263-929-30
: Sequence 30, Application US/10263929
: GENERAL INFORMATION:
: APPLICANT: Kim, Jaeseob
: APPLICANT: Galant, Ron
: TITLE OF INVENTION: Alzheimer's Disease Linked Genes
: FILE REFERENCE: USD-07417
: CURRENT APPLICATION NUMBER: US/10/263,929
: NUMBER OF SEQ ID NOS: 213
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 30
: LENGTH: 4802
: TYPE: DNA
: ORGANISM: Drosophila melanogaster
US-10-263-929-30

Alignment Scores:
Pred. No.: 0.119 Length: 4802
Score: 148.50 Matches: 190
Percent Similarity: 33.50% Conservative: 146
Best Local Similarity: 18.94% Mismatches: 376
Query Match: 3.14% Indels: 291
DB: 9 Caps: 48

US-09-914-168-2 (1-919) x US-10-263-929-30 (1-4802)
Oy 8 AlaAsnArgSerPheMetProValAlaLeuAlaAlaTyrLeuProLeuMetThrSerGln 27
Db 1867 GCAACAGAGCTTACAAATCTCGTCCCAACACAGCAGTTCTGGACCTGACCAACACAG 1926
Oy 28 ---AlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisAsp 46
Db 1927 TACAGATTAAGTACCAGCAGCAATCTGCTCCATG-----AAGTTAAGCCCTGGCGC 1980
Oy 47 ThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGlnGlnIle 66
Db 1981 AGCACACTTATAGAAATCG-----CTGTTTGAATGGCTCGAGAGAGTTT 2025
Oy 67 GlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeu----- 84
Db 2026 GATGCT-----TCTGTAGAGCGCTTCAATCTGAAGCCCAAGTGAACGATTAGTTATC 2079
Oy 85 -----AspValValAsnPheAspAspGlnSerProIleSerArgIleGlyGlu----- 100
Db 2080 AAACCCAAAGAGTAAAGTGAAGGTGGAATCCCTCTACTACTATTGAGTGGCTCC 2139
Oy 101 -----GlnSerProProLeuGlyLeuAspMetSerValIleGluGluThrThrPro 117
```

```
Db 2140 AACACGCCCAATCTGCAGCCAAAGCC-----GCAACTCC 2175
Oy 118 LeuSerLeuGluGluLeuPheAla----- 125
Db 2176 AACAAAGAACCGAGTCTTTAGCGAGCCATTCAGTAGACCGGTACACACAGCTGGA 2235
Oy 126 -----GlnGluSerThrGluMetGly----- 132
Db 2236 AACTCGCTGGAGCAACGAAATGAGCGAAAGTCAAGATACGCGCCGCTGAGTCATGC 2295
Oy 133 IleAsnProAsn-----AspTyrIleProGluTyr-----Gln 143
Db 2296 CTGATCTCTTAACAACTTTGAAGAGTACGCCAGCAACATTCAGATTCGATGATGATCAG 2335
Oy 144 GlyGluGlnProAsnSer-----GluValValValProProThrLeuGluProGlu 160
Db 2356 GGATCGCCCATTAACAGACACCTCAACAGAGCTTGCCACGTTAAACCTTTAGATACATAC 2415
Oy 161 LysProGlyLeuIleLysArgLeuTyrAlaArgLeu-----PheAsnAsp 175
Db 2416 CGACCATCTCCACCGCTTCGCTCGGTCAACCATCCAGAAATCCATTGTGAGGAT 2475
Oy 176 GlyValAsnLysValProArgLeuLysValLysPheTyrGlnSerGlnSerGlyGlu 195
Db 2476 CAATCCACACCATTTGCACGTCGCGAACC-----TTCATCTGCACAGCGCCAAATGAG 2529
Oy 196 ThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsn---IleLysAla 214
Db 2530 AGTGCTCTTTCGAAACAGATCTTAATAGGACAGAGACTCTGCAGCAACCAATCCCGCTTA 2589
Oy 215 AlaLeuGluAspIleThrGlnGlnSerAlaMetLysPheLysGlnGlySerIleProArgLeu 234
Db 2590 GCCATTGAGCGAGCTGCTGCAGGACAGACAGACAC-----GAATCA 2631
Oy 235 ArgGlnThrAlaLeuValAlaAlaArgAlaGlyTyrTyrTyrAspIleAspLeuSerIle 254
Db 2632 CATCCACACGAGCATTTG---CTGCGAGCGCTGGGTTACTTACAAATTTCA----- 2679
Oy 255 IleArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGlyGluProValTyr 274
Db 2680 -----TCGCTAAGACAGCTACGCTCTATCTGGCGGAGAGCATCTGTGGT 2730
Oy 275 IleAspTyrArgAlaValAlaGluValArgGlyGlnGlyAlaAspArgLysAlaPheThr 294
Db 2731 CCGAAGCTTT-----ACGCTGGTGGGAGAGATACGCAATGCTTCTTGGCGAAG 2781
Oy 295 ValAlaAspGluValProLeuLeuIleGlyAspValPheHisGlnGlyTyrGluThr 314
Db 2782 GAATGGATGTGGCTGGCTGACCTGATGATGATGATGATGATGATGATGATGATGATGAT 2841
Oy 315 LysLysAsnLeuIleGluAsnAlaSerAlaGlnHisGlyTyrPheAspGlyArgTyrLeu 334
Db 2841 ----- 2841
Oy 335 AspArgSerValAspValIleLeuProAsp-----AsnThrAlaAspValSerLeuIle 352
Db 2842 -----ATCATTTATCCGAGCAGCAAAACAAAGCCGCCATTTGTCAAAGT 2886
Oy 353 TyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIleAspProLys 372
Db 2887 CTAAACGCTGACGCCGAGGTGACTTTGGACCAATA-----TGGCGTTGGACAAGACC 2940
Oy 373 ThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGluGln 392
Db 2941 AAGCAGCAAGCCATCAGAGCCACACAGCTTG-----*----- 2973
Oy 393 LeuLeuThrValAsnMetGlyGluAlaTyrAsnLeuGlnAlaValAlaGlyAlaLeuSerAsn 412
Db 2974 -----CTTGAGATGTGAGCTGGAGGCGCAAGTTGGCGCGCGG-----TGGGATAAAAAC 3021
Oy 413 AspLeuIleAlaThrArgTyr-----PheAsnMetValAsnThrGluIleValIlePhe--- 429
Db 3022 GAT-----ACGCGCTTTATAGAAATACCGCCCGGAACCGGAGTGGGTCTTCCGC 3072
```

QY	430	-----	ProGluArgGluGlnIleGlnAsn	437
Db	3073	GTCAAGCACTTTTCTAATAATATGGCTGGAGACAGCGACGACAAAGAGACGGATGGCCACT	3132	
QY	438	AspGlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGlu	457	
Db	3133	GACCGGAAAAAAGCCAAATAATGGCAAGCTTGAGGCTCAGCAGGGGGCAACGGACGAAAG	3192	
QY	458	SerThrLeuGluProValIleGluThrValGluLeuThrAsp-----	471	
Db	3193	ATGACGCTCAATTCGCTGGCGCCAGAGCGCAGAAATTTTCAACAAATGACAGCCGTAACCTG	3252	
QY	472	-----	GlyIleLeuMetAspIleSerProIleGluPheSerAla	484
Db	3253	GACCCAAAAGCCTTGGCTGCCAGGTGTGGCCAGCGGATTCAGGCCAATGATGATACCGCC	3312	
QY	485	SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArg-----His	501	
Db	3313	GAGTTCCTTTTATATGACAAACACAAATTTTTCAGACCTGGTGGCACTCCGATTTCTCC	3372	
QY	502	LeuTyrAspMetProAspAspArgValLeuAlaIleAsnHisAspAspGlyValAsnArg	521	
Db	3373	ATGTTTGATTCACCGCGGTCAACGCTCCACAGATTACCAAGTCCACAGCTGTTTGGCTCAG	3432	
QY	522	SerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAsp	541	
Db	3433	GAATGGTAGGCAACGAGGGGCAACAAATGACAGTATGAAATGATATCTTTGTGGAG	3492	
QY	542	GluSerGluAsnGluValIleAspLeuProGluArgThrAla-----LeuAlaAsn	558	
Db	3493	GACAAAGTCTCCGAG-----GACGAACCCATGAGACTGAGGGGCGCTCTTTCGTGAT	3546	
QY	559	ArgLys-----	Thr	561
Db	3547	CCCAAGTTTTCACAGCGAACCCTCTTGTGGAAAGATGGTGGCTCCGAGAGTTCAGT	3606	
QY	562	ProAlaAspVal---TyrGlnSerLysValProLeuTyrValPheValAlaSer---	579	
Db	3607	CAGTACGATTTTGAAACACCATGGCCGCTCCGCGATTTCTTCCTCCGCTCGAGGCC	3666	
QY	580	-----AspLysProArgAspGlyGlnIleGlyLeuGlyTrpGlySerAspThr	595	
Db	3667	AGCGTATGTGTATGTCACACTTATGAGACACATCGTCGATGGCACTGGACAGATCTGT	3726	
QY	596	GlyThrArgLeuValThrLysPheGlnHisAsnLeuIleAsnArgAspGlyTyrGlnAla	615	
Db	3727	CGAGCTGTAAAGAAAGAAAGAAATTTGAAATCCCTGTCCTAAACCCCTCAAGTTGCTGT	3786	
QY	616	GlyAlaGluLeuLeuArgLeuSerGluAspLysGlyValLysLeuTyrAlaThr-----	633	
Db	3787	AAGCCAAAGTG-----GCTCCGATTAAAGTGCTGTGCCACACGCTT	3828	
QY	634	-----LysProLeuSerHisProLeuAsnAspGlnIleuArgAlaThrLeuGlyTyr	650	
Db	3829	CCCGTACCAGGCTCATCGCTTATGAGATCGGAGATATGATGCGCATTTGGATTC	3888	
QY	651	GlnGlnGluValPheGlnHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu	670	
Db	3889	-----TCAAAAGCAGCAGCGCTTTAAGCTCAGTTTGGTCA-----	3924	
QY	671	HisGluIleSerArgSerIleIleGlnAsnGlyLysTrpAsnArgThrTyrSerLeuArg	690	
Db	3925	-----CAAAACATCTTGGTACTGCCAAGCASCATATATCAATGACAAACCTAAA	3975	
QY	691	TyrArgLeuAspLysLeuTyrThrGlnAlaProGluIuThrTrpGlnAspLeuProVal	710	
Db	3976	-----GAATTACTGGACATCCCTCCGCTC	4002	
QY	711	AspPheValAsnGlyLysProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLys	730	
Db	4003	TCCATGATTTTTCGGCCGCTCGGCCACGATTTG-----TCGCCCAAGTATATCAG	4056	

```

0Y 731 ThrValAlaIAspAsnLeuValAsnProMetAlaArgGlyTyrArgGlnAlaTyr-----Ser 748
    |||
    |||
Db 4057 TTGGGTAGAGTTTAAACATGAGCTCAAAAGGCACACGAGGGTTTTGCGAAAGATATCATTTCCCAT 4116
    |||
0Y 749 LeuGlnValGlySerSerGlyLeuValSer----- 758
    |||
    |||
Db 4117 CTCGAACTTCAGGCTTAAACGACGCTGCTCATGTTAAATGTGGAGGGATGAGTGGCCGTGT 4176
    |||
0Y 759 -----AspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPhe 775
    |||
    |||
Db 4177 ATACATCCCGCATACGGCTAGCAAACTAGTT----- 4206
    |||
0Y 776 GlyAspAsnAlaTyrGlySerAsnAlaArgAlaHisGlnMetThrGlyGlyIleGlnAlaGly 795
    |||
    |||
Db 4207 -----TCCAAAGCATTTCTCCGAGTCTTTAAACACAGAAATGCTGGCGCTTAAAGAGCAG 4260
    |||
0Y 796 Tyr-----IleTrpSerAspAsnPheAsnHisValProTyrArgLeuAlaArgPhe 812
    |||
    |||
Db 4261 TATACCGCTGCATGCTGTGTCG-----CTTTTATTTCCCA 4293
    |||
0Y 813 AlaGlyGlyAspGlnSerIleAlaArgGlyTyrAlaHisAspSerLeuSerProIleSerAsp 832
    |||
    |||
Db 4294 CTATGGGGCGCAT-----CACAGTAAAGCTGTGAGACTGGAAAG 4332
    |||
0Y 833 LysGlyGlyTyr----- 835
    |||
Db 4333 AACTCTCATTTACATGATGATGTGCGCGCCCATCTGATCCAGATGCGCTTGAGAACAGC 4392
    |||
0Y 836 LeuThrGlyGlyGlnValLeuAlaValAlaGlyThrAlaGlyTyrAsnTyr---GluPheMet 854
    |||
    |||
Db 4393 CTGCTGGGCAAAAGATCTTACTCTGTGAAAAAAGTATGACACCCACACGCTACTGTGAGCACATG 4452
    |||
0Y 855 LysAspLeu 857
    |||
Db 4453 TTGGATTTTA 4461
    |||
RESULT 35
US-10-267-502-24
; Sequence 24, Application US/10267502
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: USD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ. ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 24
; LENGTH: 4802
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-10-267-502-24

Alignment Scores:
Pred. No.: 0.119 Length: 4802
Score: 148.50 Matches: 190
Percent Similarity: 33.50% Conservative: 146
Best Local Similarity: 18.94% Mismatches: 376
Query Match: 3.14% Indels: 291
DB: 9 Gaps: 48

US-09-914-168-2 (1-919) x US-10-267-502-24 (1-4802)
0Y 8 AlaAsnArgSerPheMetProValAlaIleAlaLeuAlaIleTyrLeuProLeuMetThrSerGln 27
    |||
    |||
Db 1867 GCACACAAAGCTTACAAATCTCTGCTGCCAACAAGCAGTTTGTGACCTGACCAAGCAACAG 1926
    |||
0Y 28 ---AlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValProLeuIleHisAsp 46
    |||
    |||
Db 1927 TACAAAGTAACTATACCACACACATCTCGCTCCCAATG-----AAGTAAAGCCGCTGGGC 1980
    |||
0Y 47 ThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGlnIle 66

```

```

Db 1981 AGCACACTTATAGAAAATCG-----CTGTTTGATGCTCTGACAGACTTT 2025
Oy 67 GlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeu----- 84
Db 2026 GATGCT-----TCTGAGAGGGCTTCAATCTGAGCCCGCATTAACGATTAGTATC 2079
Oy 85 -----AspValValAsnPheAspAspGlnSerProIleSerArgIleGlyGlu----- 100
Db 2080 AAACCCAAAGTAAAAAGTGAAGGTGAAGAAATCCCTCTAGCTCTATTGGCAGTCC 2139
Oy 101 -----GlnSerProProLeuGlyLeuAspMetSerValIleGluGluThrPro 117
Db 2140 AACACGCCCAATCTGACCCCAAGGC-----GCAACATCC 2175
Oy 118 LeuSerLeuGluLeuPheAla----- 125
Db 2176 AACAAAGACCGAGGTGTTAGCGGAGCCATTCCAGTAGAGCGCTACCACAGCTGA 2235
Oy 126 -----GlnGluSerThrGluMetGly----- 132
Db 2236 AACTGCCCTGAGCAACGAATGAGGGGAAAGTCAAGATACGCGCCGCTGACTATGG 2295
Oy 133 IleAsnProAsn-----AspTyrIleProGluTyr-----Gln 143
Db 2296 CTGATCTTAACCACTTGAAGACGTACGCCACCAACATTGCACTGGAATGATGATCAG 2355
Oy 144 GlyGluGlnProAsnSer-----GluValValAlaProProThrLeuGluProGlu 160
Db 2356 GCATCGCCCATTAACACACACCTCAAGACGCTTGCCACAGTAACCTTGAATACATAC 2415
Oy 161 LysProGlyLeuIleLysArgLeuTyrAlaArgLeu-----PheAsnAsp 175
Db 2416 CGACACATCTCCACCGCTTCCGCTCTCGGTCAACCATCCACAGAAATTCATTGAGAGT 2475
Oy 176 GlyValAsnLysValProArgLeuLysAlaLysPheThrGlnSerGlnSerGlyGlu 195
Db 2476 CAATCCAGCACCATTCGACGCTCGGAAAC-----TTCACTTCGACGACGAGCCATGAG 2539
Oy 196 ThrSerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsn-----IleLysAla 214
Db 2530 AGTGTCTTTCGAACACATCTAATGAGCGACAGACTCTGCAGCAACCAATCCCGCTA 2589
Oy 215 AlaLeuGlnAspIleThrGlnGlnSerAlaMetLysPheLysnGlnSerIleProArgLeu 234
Db 2590 GCCATTGACGACCTGCTGACAGGCGACGACGAC-----GAATCA 2631
Oy 235 ArgGlnThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspLeuSerIle 254
Db 2632 CATCCAAACAGGCAATG-----CTCGACGCGCTGCTACTATACATTTCA----- 2679
Oy 255 IleArgAsnSerIleGlyGluValAspValIleIleHisAspLeuGlyLysProValTyr 274
Db 2680 -----TCCGTAGACGACCTACGCTCTATCTGCGCGAGCGAGATCCTGTGTG 2730
Oy 275 IleAspTyrArgAlaValIleGluValArgGlyGluGlyAlaAspLysAlaPheThr 294
Db 2731 CCGCAACTT-----ACGGTGGGTGCGGAAGGATACCGCAATGCTTCTTGGCGCAAG 2781
Oy 295 ValAlaAspGluValProLeuLeuIleGlyAspValPheHisHisGlyLysTyrGluThr 314
Db 2782 GAATGATGTGGCTGGCTGCAACCTCGATGACATTTCTCCACTTTCGCAACAGACATTT 2841
Oy 315 LysLysAsnLeuIleGluAsnAlaSerAlaGluHisGlyTyrPheAspGlyArgTyrLeu 334
Db 2841 ----- 2841
Oy 335 AspArgSerValAspValIleLeuProAsp-----AsnThrAlaAspValSerLeuIle 352
Db 2842 -----ATCATTTATCCGACGACGAAACAGCCCAATGCTGCTCAAGCT 2886
Oy 353 TyrAspThrGlyThrGlnTyrArgPheAspGluValValPhePheThrIleAspProLys 372
Db -----TACAAAGACGACGCTTTAAGCTCAAGTTTGGTCCA----- 3924

```

```

Db 2887 CTAAACCGTGACGCCGAGCTCACTTTTGACCAGTA-----TGGCCGTTGACAAAGACC 2940
Oy 373 ThrAsnGlnLeuThrThrAspProAspLysLeuProValLysArgGluLeuGluGln 392
Db 2941 AAGCAGACAGCCATCAAGACCCACAAAGCTTG----- 2973
Oy 393 LeuLeuThrValAsnMetGlyAlaIleTyrAsnLeuGlnAlaValAlaArgAlaLeuSerAsn 412
Db 2974 -----CTTGAGATGAGACTGCGAGGCAAGTTGGCGCGGTG-----TGCATAAAAAC 3021
Oy 413 AspLeuIleAlaThrArgTyr-----PheAsnMetValAsnThrGluIleValPhe--- 429
Db 3022 GAT-----ACGCGCTTTATGAATACCGCCGGAACCGGTACTGGCTTCCGC 3072
Oy 430 -----ProGluArgGluGlnIleGlnAsn 437
Db 3073 GTCAGACATTTCTAATATGCGCTGGAGACAGCAGCAAGAGACGAGTTCCACT 3132
Oy 438 AspGlnValSerPheGluGlnSerSerSerArgThrGluProAlaGlnValAspGlu 457
Db 3133 GACCCGAAAAAGCCCAAAATTTGCAACGCTTGAGGCTCAGACGCGCAACGCAAGAAAG 3192
Oy 458 SerThrLeuGluProValIleGluThrValGluLeuThrAsp----- 471
Db 3193 ATGACGCTCAATTGCTGCGCCACGCGCAGAGAAATTTCAAGATGACGCGCTAAGCTTG 3252
Oy 472 -----GlyIleLeuMetAspIleSerProIleGluPheSerAla 484
Db 3253 GACCCAAAGCCTTGTCGCGAGGCTTGCCAGCGAGATTACGCCAATGATGATACGCC 3312
Oy 485 SerAsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArg-----His 501
Db 3313 GAGTTCCTTTAATGAGCAAAACCAATTTTCCAACTGCTGTCGCACTCCGATTTCCTC 3372
Oy 502 LeuTyrAspMetProAspArgValLeuAlaIleAsnHisAspArgLysAlaAspArg 521
Db 3373 ATGTTTATCCACCGCGCTCAACGCTCAACGATTACAGCTCCACAGCTGTTTGCTCAG 3432
Oy 522 SerIleLeuGlyArgIleSerAspAlaValSerAlaValAlaArgAlaIleLeuProAsp 541
Db 3433 GAATGAGTAGCAACGAGCGACCAAAATCCAGCTAATGAAGTCATCTTGTGAG 3492
Oy 542 GluSerGlnAsnGluValIleAspLeuProGluArgThrAla-----LeuAlaAsn 558
Db 3493 GACAACTCTCCGAG-----GACGAACCATGAGACTACGCGGCGCTTCTTCTGAT 3546
Oy 559 ArgLys----- 561
Db 3547 CGCAAGCTTTTCAACGTGAACCTCTGTTGGAAGAATGCTGCTCCGACAGTTCCAGT 3606
Oy 562 ProAlaAspVal---TyrGlnSerLysLysValProLeuTyrValPheValAlaSer--- 579
Db 3607 CAGTACGATTTTGAAACACCCATCGCCGCTGCGATTTCTTCCCTGCTGAGAGCC 3666
Oy 580 -----AspLysProArgAspGlyGlnIleLeuGlyLeuGlyTyrPheLysAspThr 595
Db 3667 AGCCTGATGTGATGACATTAAGAGACATCGTGCATGCACTGCAAGCATTTGTT 3726
Oy 596 GlyThrArgLeuValThrLysPheGluHisAsnLeuIleAsnArgAspGlyTyrGlnAla 615
Db 3727 GCAGCTTAAGAAACGAAGATTGAATGCTGCTCACTCAAGCTTCAAGTTTGTCTGT 3786
Oy 616 GlyAlaGluLeuArgLeuSerGluAspLysGlyValLysLeuTyrAlaThr----- 633
Db 3787 AAGCCCAAGTG-----GCTCCGATTAAAGTTGGCTGCCACAGCTT 3828
Oy 634 -----LysProLeuSerHisProLeuAsnAspGlnLeuArgAlaThrLeuGlyTyr 650
Db 3829 CCGCTACCCAGGTCACATCGCTTATGAGATCGCGGATTAATGATTCGCGGATTTGGGATTC 3888
Oy 651 GlnGlnGluValPheGlyHisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlu 670
Db -----TACAAAGACGACGCTTTAAGCTCAAGTTTGGTCCA----- 3924

```

```

QY 671 HisGluIleSerArgSerIleIleGlnAsnGlyIleTyrPasnArgThrTyrSerLeuArg 690
DB 3925 -----CAAACTCATGTGTACTGCCACAGACACTATAATCAACTGAAACCTGAAA 3975
QY 691 TyrArgLeuAspLysLeuLysThrGlnAlaProProGluThrTyrPglAsnLeuProVal 710
DB 3976 -----GAATTTACTGAGACATCCCTGCCCGCTC 4002
QY 711 AspPheValAsnGlyLysProSerGlnGlnAlaLeuLeuAlaGlyValAlaValHisLys 730
DB 4003 TCCATGATTTTGGCCGCCGCTCGGCCACCATTTG-----TCGCCAGTGAATGCGAG 4056
QY 731 ThrValAlaAspAsnLeuValAsnProMetArgGlyTyrArgGlnArgTyr-----Ser 748
DB 4057 TTGGTGGAGTTTAACATGTGTCAAGGACAGAGGGTTTGCAGAAAGTATCATTTCCCAT 4116
QY 749 LeuGluValGlySerSerGlyLeuValSer----- 758
DB 4117 CTCGAGATTACACTTAACGACTGCCTCTCACTTAATGTGAGAGGTAGTAGTCCCGTGT 4176
QY 759 -----AspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyrSerPhe 775
DB 4177 ATACATCCCGATAGCGGTGCGAACTAGT----- 4206
QY 776 GlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyIleGlnAlaGly 795
DB 4207 -----TCCAGCAGCTTCTCCAGTCTCTTAAGCAGAGAAATGTGCTTAAGAAAGAGAC 4260
QY 796 Tyr-----IleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePhe 812
DB 4261 TATACCGTGTCACTGTGGTGG-----CTTTATTTCGCA 4293
QY 813 AlaGlyLysAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAsp 832
DB 4294 CTATGCGGCGAT-----CACGATGAAGTGTGAGCACTGGAAG 4332
QY 833 LysGlyTyr----- 835
DB 4333 AACTCTCATTAATGATGTCATGTGTGGCCCAATCTGCTATCCGAGTGGCTTGAGAACAGC 4392
QY 836 LeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGlyTyrAsnTyr---GluPheMet 854
DB 4393 CGTGTGGGCAAGAAAGTACTCTCTGAAAAAGTTAGCACACAGCTGAGGACATG 4452
QY 855 LysAspLeu 857
DB 4453 TTGGATTTA 4461

RESULT 36
US-10-172-502-1
: Sequence 1, Application US/10172502
: GENERAL INFORMATION:
: APPLICANT: FOSTER, Timothy et al.
: TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.
: FILE REFERENCE: P07263US01/BAS
: CURRENT APPLICATION NUMBER: US/10/172,502
: PRIOR FILING DATE: 2002-06-17
: PRIOR APPLICATION NUMBER: US 60/298,098
: NUMBER OF SEQ ID NOS: 29
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 6609
: TYPE: DNA
: ORGANISM: Staphylococcus epidermidis
US-10-172-502-1

Alignment Scores:
Pred. No.: 0.288 Length: 6609
Score: 146.00 Matches: 185
Percent Similarity: 34.148 Conservative: 138
Best Local Similarity: 19.368 Mismatches: 365

```

```

Query Match: 3.09% Indels: 258
DB: 9 Gaps: 40
US-09-914-168-2 (1-919) x US-10-172-502-1 (1-6609)

QY 27 GlnAlaLeuAlaGlnGlnAsnAsnProAlaAsnIleIleAsnHisValProAlaHisAsp 46
DB 4201 CAAGCATTTAGCTAAAGAAAGAAAGAAAGCACTT----- 4233
QY 47 ThrAlaIleAsnGlnAlaLysAlaGlyAsnProProValLeuLeuThrProGluGlnIle 66
DB 4234 GCAGCTATTGACCAAGCT-----CAACG 4257
QY 67 GlnAlaArgLeuAsnAlaAlaGlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspVal 86
DB 4258 AATAGTCAGGTGATATCAACGCGCAACAATGTT-----GTATCAGCGATTAAAT 4308
QY 87 ValAsnPheAspAspGln---SerProIleSerArg-----IleGlyGlnGlnSerPro 103
DB 4309 ATTCAACCTGAAACAAATAAACCAGCAGCAGTGAATAAATCAATCAAAAAGCGAAT 4368
QY 104 ProLeuGlyLeuAspMetSerValIleGluGluThrThrProLeu-----CAAAG 4428
DB 4369 GAATTAAGTGGCCAAATTAATACAGATTAAAGAGCGACAGCAGACAAAGACAGCGCGG 4428
QY 119 -----SerLeuGluGluLeuPheAlaGlnGlnSerThrGluMetGlyIleAsnProAsn 136
DB 4429 TTAGATTAATATCAATGATTAGTGTGTAAGATGATGACAAATATC-----ACGAAT 4479
QY 137 AspTyrIleProGluTyrGlnGlnGluGlnProAsnSerGluValAlaValProProThr 156
DB 4480 GATAGCAACAATACGCAAGTTAATGACTCACACAAACGCTTGACGACATTTGCATTA 4539
QY 157 LeuGluProGluLysProGlyLeuIleLysArgLeuTyrAlaArgLeuPheAsnAspGly 176
DB 4540 GTGACCGCTGACCAT-----ATTGTTAGACAGCTGTAGA-----GATGCA 4581
QY 177 ValAsnLysValProArgLeuLysAlaLysPheTyrGlnSerSerGlnSerGlyGluThr 196
DB 4582 GTTAAAGCA-----CAATATGAGCTAAAGACGAATATGAGCAACGCAACATGCG 4635
QY 197 SerAlaIleGlySerSerHisGlnLysThrGluProTyrAlaAsnIleLysAlaAlaLeu 216
DB 4636 ACTGATGAAGAAACAAACAGTCTTTAAATCAATTAAGCAATATGAAAGAGCTGATTA 4695
QY 217 GluAspIleThrGlnGlnSerAlaMetAspLeuAsnGlySerIleProArgLeuArgGln 236
DB 4696 CAAACATTAAATCAAGCAATACG-----AATATGATGTGAACGTGTGAATCA 4746
QY 237 ThrAlaLeuValAlaAlaArgAlaValGlyTyrTyrAspIleAspSerIleIleArg 256
DB 4747 AATGATATTGCTACGTTAAAGCGCTA----- 4773
QY 257 AsnSerIleGlyLysValAspValIleIleHisAspLeuGlyGluPro---ValTyrIle 275
DB 4774 -----GAACCGCACATTTGGT 4791
QY 276 AspTyrArgAlaValGluValArgGlyGluGlyAlaAspLys----- 290
DB 4792 AAACCTGAGCTCAAGCAACCATAAAGCGACGAGATACCAATCAATCTATAA 4851
QY 291 -----AlaPheThrValAlaAspGluValProLeuLeuIleGlyAspVal 306
DB 4852 GATACACCAATGCTAGCAGACATGATTAATGAAGCAACCAACAATTAACGACACA 4911
QY 307 PheHisHisGlyLysTyrGluThrLysLysAsnLeuIleGlnAsnAlaSerAlaGluHis 326
DB 4912 CTTAAACAAGGTCAACAAGAT-----ATGACAAATTCGACAAAGATGCA 4956
QY 327 GlyTyrPheAspGlyArg---TrpLeuAspArgSerValAlaIleLeuPro----- 343
DB 4957 GCTGTCAATGATGTGAAGAACCAACGAATTAAGGCAATGCAATTAATTAACCGAAAGTT 5016

```

```

OY 344 -----AspansThralaAspValSerLeuIleTyrAspThrGly 356
DB 5017 AGACGCAAACTGCAGCGTTGGATAAC----- 5043
OY 357 ThrGlnTyrArgPheAspGluValAlaPhePheThrIleAspProLysThrAsnGlnLeu 376
DB 5044 -----ATTGATGGAAGTAATAATAATCAACTGCAATCGAATGCAATGCGTA 5091
OY 377 ThrThrAspProAspLysLeuProValLysArgGluLeuGluGlnLeuLeuThrVal 396
DB 5092 GATACACCGCAAGATGCAAGAAATGCTGATTCGCGTTAAATAAATT----- 5142
OY 397 AsnMetGlyGluAlaTyrAsnLeuGlnAlaValAlaLeuSerAsnAspLeuIleAla 416
DB 5143 -----GTTAATGCAATTGATAAATAAATGATTCACAA 5172
OY 417 ThrArgTyrPheAsnMetValAsnThrGluIleValPheProGluAlaArgGlnIleGln 436
DB 5173 -----ACNAACGCAATCCAGAAAGTGATCAAACTGAGCGTGATGTAACAC 5220
OY 437 AsnAspGlnVal-----SerPheGluGlnSerSerSer 447
DB 5221 AACATCAAAAGGATTTTACCTAAAGTTAACGACGAGCGCGTCATCTGTCAGC 5280
OY 448 SerArgThrGluPro-----AlaGlnValAspGluSerThrLeuGluProValIleGlu 465
DB 5281 GCAAAAGCTGAAGCTCAAAATGCTACTTATGATCAAGTGAATTTATACCGAAGAAAGAA 5340
OY 466 ThrValGluLeuThrAspGlyIleuMetAspIleSerProIleGluPheSerLaser 485
DB 5341 -----AGATTAGCTGCTTAA 5355
OY 486 AsnLeuIleGlnAspLysLeuAsnLeuValAlaAlaLysAlaArgHisLeuTyrAspMet 505
DB 5356 CATTTAGTAGAACAACGACTTAATCAACCTATTGATCAAGCAATACACGCA----- 5406
OY 506 ProAspAspArgValIleuAlaIleAsnHisAspAspGlyValAsnArgSerIleLeuGly 525
DB 5407 -----GATTAGACTCGCCCAACTTATCAAAATAGTATCGCTCAAAATAATTTTCA 5460
OY 526 ArgIleSerAspAlaValSerAlaValAlaArgAlaIle-----LeuProAspGluSer 543
DB 5461 AAAATTAACCGACGACACAGTTAAACGACATTAACAACAAATTCAAATATATCGCT 5520
OY 544 GluAsnGluValIleAspLeuProGluArgThrAlaLeuAlaAsnArgLysThrProAla 563
DB 5521 ACAATTAATAATTAATTAATAAGCAATACGACGACGATGACGACGACGACGACGACG 5580
OY 564 AspValTyrGlnSerLysLysValProLeuTyrValPheValAlaSerAspLysProArg 583
DB 5581 GCATTAAGTACAAAGTTGAAAGAGCTAAT-----AAAGCTTAA 5619
OY 584 AspGlyGlnIleGlyLeuGlyTyrPGLySerAspThrGlyThrArgLeuValThrLysPhe 603
DB 5620 CACCAAAATTCGCTGCT-----GCAGTACTAATGCT 5649
OY 604 GluHisAsnLeuIleAsnArgAspGlyTyrGlnAlaGlyAlaGluLeuLeuSerGlu 623
DB 5650 GATGTCGATATTTATTCATGATGAGGAAAAAC-----GAAATTCGGAATCGAA 5700
OY 624 AspLysLysGlyValLysLeuTyrAlaThrLysProLeuSerHisProLeuAspGln 643
DB 5701 CCTGTTATTAATAAAGCAACGACCGCCGACGACAAATTAACACATTAATTCACGATAG 5760
OY 644 -----LeuArgAlaThrLeuGlyTyrGlnGlnGlnValPheGly 656
DB 5761 AAACCAACCAATTGAAGCAATGTTCAAGCAACGTAAGCAAAAGCAATGATTTTGA 5820
OY 657 HisSerThrAsnGlyPheAspLeuSerThrArgThrLeuGlnHisGluIleSerArgSer 676
DB 5821 CAGTTACAAAACATTTATGACTGCTATTGACAAATTTATCAAGATCGTAGCAATGCA 5880
OY 677 IleIleGlnAsnGlyLysTyrPasnArgThrTyrSerLeuArgTyrArgLeuAspLysLeu 696

```

```

DB 5881 -----CAAGTTCATTAACCA 5895
OY 637 LysThrGlnAlaProGluThrTyrPGLAspLeuProValAspPheValAsnGlyLys 716
DB 5896 GCAACA-----TTAAATCTCAAAACATACATGATTTAGACGTACATCCATTAACAAAG 5949
OY 717 ProSerGlnGluAlaLeuLeuAlaGlyValAlaValHisLysThrValAlaAspAsnLeu 736
DB 5950 CCAGATGCTGAA-----AAACGATTAATGATGATCTT 5982
OY 737 -----ValAsnProMetArgGlyTyrArgGlnArgTyrSerLeuGluValGlySer 753
DB 5983 GCACGCTTTACACATTTAGTCGCAAAATATGCAAA----- 6018
OY 754 SerGlyLeuValSerAspAlaAsnMetAlaIleAlaArgAlaGlyIleSerGlyValTyr 773
DB 6019 -----GTAAGTGAATGTAATTAAGCTGATGCTATTAAGCTATTAACGATTTAA 6069
OY 774 SerPheGlyAspAsnAlaTyrGlySerAsnArgAlaHisGlnMetThrGlyGlyIleGln 793
DB 6070 TTACAAATGATGAAGATTTAAAAACACACGCTAATGCT----- 6111
OY 794 AlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArgLeuArgPhePheAla 813
DB 6112 -----GATGTTGATGCAGTTTAAACGATTTAAATGTTGCAATTA 6150
OY 814 GlyLysAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSerProIleSerAspLys 833
DB 6151 GGCAGTTATAGAACAGATTAATTAAGTAAAGAAAGAAAGAAAGTAAAGTAAAGTAAAGTAA 6210
OY 834 GlyTyrLeuThrGlyGlyGlnValLeuAlaValGlyThrAlaGluTyrAsnTyrGluPhe 853
DB 6211 GCTCAACAAACATATGCGAAATTCAAAGCGATCCGACACAGAA-----CAATTAGCTTAA 6267
OY 854 MetLysAspLeuArgLeuAlaValPheGlyAspIleGlyAsnAlaTyrAspLysGlyPhe 873
DB 6268 GTAAAGCATTAATTAATGATCAATATGTTGACAGATGCAATAGCAATAGTGAAGATGCG 6327
OY 874 Thr-----AsnAspThrLys-----IleGly 880
DB 6328 ACATTAATGACATCAAAAGATACGCAACTCTTATGATGAATTTTATAGCAATTTAA 6387
OY 881 AlaGlyValGlyValArgTyrPalaSerPro-----ValGlyGlnValArgValAspValAla 899
DB 6388 TTACCTGCTGAAGATTAAGGCGTCAACAAAGTGGGCAACCTGCTCCAAAAGTTTGT 6447
OY 900 ThrGlyValLysGluGlu 905
DB 6448 ACGCCTATTAAAAAGAA 6465

RESULT 37
US-10-320-800-39
: Sequence 39, Application US/10320800
: GENERAL INFORMATION:
: APPLICANT: ROBINSON, ANDREW
: APPLICANT: GORRINGE, ANDREW
: APPLICANT: HODSON, MICHAEL
: APPLICANT: REDDIN, KAREN
: TITLE OF INVENTION: MULTICOMPONENT MENINGOCOCCAL VACCINE
: FILE REFERENCE: 1581.0790001
: CURRENT APPLICATION NUMBER: US/10/320,800
: PRIOR FILING DATE: 2002-12-17
: PRIOR APPLICATION NUMBER: PCT/GB99/03626
: NUMBER OF SEQ ID NOS: 75
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 39
: LENGTH: 7542
: TYPE: DNA
: ORGANISM: Neisseria meningitidis
: FEATURE:
: NAME/KEY: CDS

```


Oy	585	yglHilEclYleuLlyrprcIysErSerPrrGlyThrArgLeuValThlYsrPhgIIHI	605
Db	2499	CGArTGTGGCGTTGGCGGCCCATTAAGTATACCAATGGCCAAAGCCACAGCCAAAGGCCAAAG	2558
Oy	605	sasNleuLIeasNrgraspGly-----TyrgInalagIyalagIuleuArgrLeuSergI	623
Db	2559	G-----ATACACAMTCCCGCCCAACCATTCAGCTCGAGCCAAACCTGTTTAGTGCT	2612
Oy	623	wAsPrLyelYslYalYlsYleuYrAlathrlYsProLeuSerHIsProLeuAsnAsgI	643
Db	2613	AGAAAAG-----CTGCACAATACCATAATGA	2639
Oy	643	nLeuArgrAlathrlEu-----glyTYrgInGln-----GlUvalPheGI	656
Db	2640	TTrtGAAACACGATTGGGTAGAACAGGCGCGAGCATATTGTGTGAATTACGAAGCATTTGG	2699
Oy	656	yHISserThrasnClYsheaspleuSerThrargThrlengluIHIsGluIlleserArge	676
Db	2700	ACGA-----CAGCAATTTATTTGGAGAGGACGACGAAATGAATTAAGCTGGTC	2747
Oy	676	rIleIleGln-----AsnGLyGlyTr	683
Db	2748	TGTCTATACGATGATCATCACACACTTACGACCCCTGATGAGCGGCCCATGAATAATG	2807
Oy	683	pAsnarGrThrlYsrleuArgrYrArgLeuasplYsleu-----LysThrGln-----	699
Db	2808	GCAT-----AAATACCATTTATGAAAAAGTCACCCAAACCAACCCAAAGTTAC	2852
Oy	700	-----Alapro-ProGluThrTrpGlnAspLeuroyAlasrPhevalasnglYsLys-	716
Db	2853	CCAACGTCCGCGACCAAATCATCTTCAGGTATGATTAAACCATTTGATGATGAAGAAGT	2912
Oy	717	-----ProserGlnGlualeuLeuAlagIyalaIavalHisLysThrValaa	734
Db	2913	ATTTAATATCCGATAGCCCAATCATCTGTGTGGCAATCTCATTTGTACAAAGAAAAGA	2972
Oy	734	sPaSnleuValaAsnrPrometArgrGLYTyrArgGlnArgrYsrLeugluValIGlysTs	754
Db	2973	CGGTTTTGCATACGACGAACCTTTGGCCAAAGAAAGATTTACGTGAATAATGGCAATT	3032
Oy	754	erGIyleuLeuLIsErspAlaAsn-----M	762
Db	3033	ACACAGCATTATGGCTGAGAAACATGAAGAGAGACTCAACGGGACATAGCGAACMAA	3092
Oy	762	eTalialIlealArgralagIylIleserGlyalYsrPheGlysPasnalArtyrGls	782
Db	3093	TTrACATTTTCCCGGAGGAA-ATACACGCGACATTTCTACTGGGTTTCATTTGGCTATGAT	3151
Oy	782	erEsnArgr-----AlahIsGln-----MetHrG	790
Db	3152	CGCATCCCAAGACATTTAAGCCATCATTCGCCGCCCAAGACGACTGATGCCGCAAAACA	3211
Oy	790	IyglYllIecInalagIlyTrIlerPsrEsrAspnPheasnHIsValProTYrArgLeua	810
Db	3212	ACGGTATTTCGCTACCCCTATACG---TCCAAATCTTTTACCCCATTTACCACACACAGCT	3268
Oy	810	rgPhePhealagIyLyspClnserLlearGlyYrAlahIspsrLeuSerProI	830
Db	3269	TATTCATT-----ATCAATCTCTG	3286
Oy	830	IesErAsplysGlyTYrLeu	836
Db	3287	TC---AATTAAGGCTATCTT	3303
RESULT 38			
US-10-282-122A-29937			
Sequence 29937, Application US/10282122A			
GENERAL INFORMATION:			
APPLICANT: Wang, Liangsu			
APPLICANT: Zamudio, Carlos			
APPLICANT: Malone, Cheryl			
APPLICANT: Haselbeck, Robert			
APPLICANT: Ohlsen, Karl			

```

APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/220,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 29937
LENGTH: 7545
TYPE: DNA
ORGANISM: Neisseria meningitidis
US-10-282-122A-29937

```

Alignment Scores:			
Pred. No.:	0.376	Length:	7545
Score:	145.50	Matches:	187
Percent Similarity:	34.81%	Conservative:	150
Best Local Similarity:	19.32%	Mismatches:	336
Query Match:	3.08%	Indels:	297
DB:	9	Gaps:	48
US-09-914-168-2 (1-919) x US-10-282-122A-29937 (1-7545)			
QY	19	AlAtyrLeuProLeuMetThrSerGlnAlaLeuAlaGlnGlnAsnAspProAlaAsnIle	38
Db	841	GCACATTTCGCTATTCTCATTAATGCTGCGCCAAATACCTCAACAAATACAGCCAAAC	900
QY	39	ILeAsnHisValProAlaHisAspThrAlaIleAsnGlnAlaIleGlyAsnPro---	57
Db	901	GGCACACATATCCCTTTATT-----GGCATTTGATACAGCCAAATTTAGGAGCTATGTAT	954
QY	58	-----ProValLeuLeuThrProGlnGlnIleGlnAlaArgLeuAsnAlaIa	73
Db	955	GCCAAACAAATACACCTTGATGCAGTACAGCCGCCAGACAGCAGCATTCGT---AATCAAGG	1011
QY	74	GlyLeuAsnAlaLysProGlnSerGlnAlaLeuAspValIValAsnPhaAspAspIns	93
Db	1012	CAGTTGGTTTCCTCTTCGCCGTAATGNGGGATGATGAC-----AATGCC	1056
QY	94	ProIleSerArgIleGlyGlnGlnIns-ProproLeuGlyLeuAspMetSerValIleG	113
Db	1057	CGTTTATCATCATAGTGGCAGACAGTGCCTGCCGCCAAG---CGAAAGATACGGATATATACA	1113
QY	113	uGluThrThrProLeuSerLeuGlnIuLeuPheAlaGlnGlnInsThrGluMetGlyI	133
Db	1114	GCGGACACACAA--GTCAATATCCG-----AGTCAGGCGGT	1148

Oy	133	easnpProAsnAspPrrTLleProSLuYrGngLuGlnpProAsnSer-----	143
Db	1149	TGAaAACGCGGTACGGCGGGTATCGCAACAGGGCACTCAAAATTCACAGTCAGTCGATTCA	1208
Oy	150	-----GluValValVal-----ProProThleuGL	158
Db	1209	AAACATCGGCACATTATTGTTCCTCAGGCGCAAAATTTGATTGCATTCATTTGGGAGCGCTGAA	1268
Oy	158	uPrgGLuLysProGLuLeuLleLysArgLeuTYrAlaArgLeuPheAsnAspGLuValAs	178
Db	1269	AAATGAAGAATCATCGACGCACATTGAACCGCGCTGTTGGCTATTGATTACCGACACACTTAA	1328
Oy	178	uLysValLProArgLeu-----LysAlaLysPheTYrGlnSerSerGln-----	192
Db	1329	TAAATCAAGCAGCAAACTCTCTCAAAACAGTTCCAAAAAATCCATATTGATGCACAAAGCAA	1388
Oy	193	-----SerGLyGLuThrSerAlaLleGLySe	201
Db	1389	AATGATTAACCGTGGCGCGCATGGTTTTACAAAGATACCGCACCAACCGGTCAAATGGTTTC	1448
Oy	201	rSerHisGLuLysThrLuproTYr-----GluValValVal-----	209
Db	1449	AAGCAATCAAAACCGGCATAGTTACATATGCATCTTCCATTATCACTACACCAACCAAC	1508
Oy	210	-----AlaAsnLleLysAla-----	214
Db	1509	AACGGCAACAGTACGGGTACTCGCAACCGTTTCTATTATCAACATATACGTGGCTTACCTT	1568
Oy	215	-AlaLeuGLuAspLleThrInGLuSerAlaMetAspLeuAsnGLySerLleProAlaLe	234
Db	1569	TGCTGATGGGACATTTGCACATGATGGTGACGTGACGTAAATTGACGACGATTAATTTGCCAA	1628
Oy	234	uArgGLuThrLalaLeuValAlaAlaIArgAlaValaLGLyTYrTYrAspLleAspLeuSerLl	254
Db	1629	TGGTCAACACAGATTTGTAGTGGCAGCAAGTTTA-----	1662
Oy	254	eIleArgAsnSerLleGLyGLuValAspValLleLleHisAspLeuGLyLuproValTY	274
Db	1663	-----AAATATCTACAGACAAATATAGAC-----ATTCTATCGTTA-----	1693
Oy	274	rLleAspTYrArgAlaValaLGLuValaArgGLyGLuGLyAlaAspAspLysAlaPheThrTh	294
Db	1696	-----AATGCCAAAGGTTGGCGGCTTGGACATTCACATCAACATGGAACAAAT	1736
Oy	294	rValAlaAspGLuValaProLeuLeuLleGLyAspValPheHis--HisGLyLysTYrGL	313
Db	1737	TATCAGTATCGGTCCATCCATTCACAGCGGCACCGCTGATTAATCAAAATGGCAACATCAAC	1766
Oy	313	uThrLysLysAsnLeu-----LleGLuAsnAlaSerAlaGLuHisGL	327
Db	1797	AACACGGCCAAACAGTTAAAGATTTGAACCCGATCAACATGGATTAACGCT-----CATGG	1847
Oy	327	y---TYrPheAspLysPrrTYrPheLysAspAspSerValaAspValLleLeuProAspAspTh	346
Db	1848	CAAGTATTATTACACAGAAATATAGCGGATTAGCCGTTTCAGAGACGCGTCAACAAATCAAAA	1907
Oy	346	rAlaAspValSer-----LeuLleTYrAspThrGLyThrGLnTYrArgPheAs	362
Db	1908	TGGCGGAATATACCGACCAATCAACACACTGATTATTCACGATGGTCAGACATCTTACC-----	1962
Oy	362	pGLuValValaPhePheThrLleAspProLysThrAsnGLnLeuThrThrAspProAspLY	382
Db	1963	-----GCTGTCAATGTGAATATACCAATGGC-----	1986
Oy	382	sLeuProValLysArgGLuLeuLeuGLnLeuLeuThrValaAsnMetGLyGLuAlaTY	402
Db	1987	-----ACGATATCAATCAACGCGCTCGATGT	2009
Oy	402	rAsnLeuGLuValaValaArgAlaLeuSerAsnAsp-----LeuLleAlaThrArgTYrPh	420
Db	2010	TGCTATTTCAGGCA-----AAATTCGTTATTCACCAACGGCGACACTTGGCGCGGTAAATAACT	2066

OY	420	eaasmeValasnthrciululevalpheprociuaargluglinileglnasaspclnva	440
Db	2067	GGAAATTTGCGCTTACAAACATATATTTTATATGATACCAATATCGCGGGGCAATGAATTT	2126
OY	440	IserPhegluGlnSerSerSerSerArgThrGluProIaGlnValaIaspGluSerThrLe	460
Db	2127	GTCCCTCACTACAGACGAGCGACCTGAAAAATTCA	2160
OY	460	ugluProValaIleGluThrValGluLeuThrAspGlyIleleuMetAspIleSerProI	480
Db	2161	-----CAATCTTTCGACACAGAAAAACGCAATTCGGATTAAAGCAAAATAACT	2207
OY	480	egluPheSerAlaSerAsnLeuIleGlnAsp	492
Db	2208	TGATATATCGACACACAGCAACATTCATATCCGCGGTACAGACAGCAATTGGCAGCCAGCA	2267
OY	492	uasnleuValaIaLeuAlaLysAlaIaArgHisLeuTyraSpmetProAspArgValaLeuI	512
Db	2268	CAATTTAAC-----AATAGAGCTTGATTGACGACACAAACCAAAATCCAAAGC	2318
OY	512	aIeasnhIsaspSpcllyValaAsnArgSerIleleuGlyArgIleSerAspAlaValSe	532
Db	2319	C-----CGGCAATGAATTAATATCGCTACAGCGTGGATTTATCGGACAAATAT	2366
OY	532	ralaValaIaIaArgAlaIaIeLeuProAspGluSerGluAsn	549
Db	2367	CGCATATGCGCGCTTACCGCTTTAACAATCAATCAATGAATAAAGCGGTACAGTCCGCCCATTCG	2426
OY	549	pleuProGluIaArgThrAlaLeuAla	565
Db	2427	GGCACGTGAAAACCTGCAATTTAGCATCGACCAATTAACCAACCGTGAACAAACATCTAT	2486
OY	565	I TyrGlnSerLysLysValaProLeuTyraIaPheValaIaSerAspLysProIaArgSpcl	585
Db	2487	TTACACGGCT	2498
OY	585	ygluIleGlyLeuGlyTyrProLysSerAspThrGlyThrArgLeuValaThrLysPheGluH	605
Db	2499	CGAATATGCGCGTTCGCGCGCATTTAGTACCAATGCGCAAGCCACAGCCAAAGCCCAAG	2558
OY	605	asnleuIleAsnArgSpcl-----TyrGlnAlaGlyAlaGluLeuArgLeuSerG	623
Db	2559	G-----ATACACAATCGCGCGCCCAACCATTTGACGTGACAGGCCAAATTCGTTTTAGGT	2612
OY	623	uasPlysLysGlyValaLysLeuTyraIaThrLysProLeuSerHisProLeuAsnAspG	643
Db	2613	AGAAAAG-----CTGCACAAATACCAATGACCA	2639
OY	643	nleuIaArgAlaThrLeu-----GlyTyrGlnGln-----GluValaPheG	656
Db	2640	TTTGAATAACCGCATTTGGTGTACAAACAGGCGCGAGCATTTGTTGATTCGAAGACATTTGG	2698
OY	656	yHisSerThrAsnLysPheAspLeuSerThrArgThrLeuGluHisGluIleSerArgSe	676
Db	2700	ACGA-----CACAAATTAATTGGCAGAAGCGACGACAATCAATTAAGCTGGTC	2747
OY	676	rIleIleGln-----AsnGlyGlyTr	683
Db	2748	TGTCTATATACGATCAATCAGACCACTTACGCCACCCCTGATGAGCGCGCCATGAAAAATTG	2807
OY	683	pasnaArgThrTySerLeuArgTyraArgLeuAspLysLeu-----LysThrGln-----	699
Db	2808	GCAT-----AAATACGATTTATGAAAAAGTCACCCCAAAAAAACCCCAAGTTAC	2852
OY	700	-----AlaPro-ProGluIaThrTyrGlnAspLeuProValaIaspPheValaSnGlyLys-	716
Db	2853	CCAAACTGCGCCACCAAAATCATTTCAAGTAATGATTAAACCATTAAGTGAAGAAAGT	2912
OY	717	-----ProSeGlnGluAlaLeuLeuGluIaGlyValaIaValaHisTySerThrValaIaA	734
Db	2913	ATTTAATATACGATGCAATCATCTGCTGTGCATATCTCATTTGTACAAACGAAAAAAGA	2972
OY	734	spasnleuValaAsnProMetArgGlyTyrArgGlnArgTySerLeuGluValaGlySerS	754

```
Db 2973 CGGTTTGCATACGACGAACCTTTGGCGAAAGAAAGTATTCAGTGAATAATGGCAATT 3032
OY 754 eRgLYLeuValSerAspAlaAsn-----M 762
Db 3093 ACACAGCATATTGGCTGAGAAACATAAAGACGACATCAACGGGACATAGCGAACAATA 3092
OY 762 eAlaIleAlaIArgAlaGlyIleSerGlyValTYrSerPheGlyAspAsnAlaTYrGlyS 782
Db 3093 TTACACTTTCCCGGAGAA-ATCACACGCAACATTTCCACTGGGTCATTTGGCTATGAAT 3151
OY 782 cTAsnArg-----AlaHisGln-----MetHisG 790
Db 3152 CGCATCGCAAGCATTAAGCCATCATCGCCGACGCAAGGACATGATGGCCCAAGCA 3211
OY 790 LYgLYleGlnAlaGlyTYrIleTYrSerAspAsnPheAsnHisValProTYrAlaGlyLeuA 810
Db 3212 ACGGATTTCCGCTACCCATAGC---TCCATTTCTTTTACCCCATTTACCCAGACAGCT 3268
OY 810 gRhePheAlaGlyGlyAspGlnSerIleArgGlyTYrAlaHisAspSerLeuSerProI 830
Db 3269 TATACATT-----ATCATCTCTG 3286
OY 830 IeserAspLYsGlyTYrLeu 836
Db 3287 TC---ATAAAGGCTATCTT 3303

RESULT 39
US-10-018-470-1/c
: Sequence 1, Application US/10018470
: GENERAL INFORMATION:
: APPLICANT: Frazer, Claire M.
: APPLICANT: Hickey, Erin
: APPLICANT: Peterson, Jeremy
: APPLICANT: Tetteilin, Herve
: APPLICANT: Venter, J. Craig
: APPLICANT: Masimanol, Vega
: APPLICANT: Galeotti, Cesira
: APPLICANT: Mora, Manrosa
: APPLICANT: Rattl, Giulio
: APPLICANT: Scarselli, Maria
: APPLICANT: Scarlato, Vincenzo
: APPLICANT: Rappuoli, Rino
: APPLICANT: Pizsa, Mariagrazia
: APPLICANT: Grandi, Guido
: TITLE OF INVENTION: Neisseria Genomic Research
: FILE REFERENCE: CHIR-0319
: CURRENT APPLICATION NUMBER: US/10/018,470
: PRIOR APPLICATION NUMBER: 60/132,068
: PRIOR FILING DATE: 1999-04-30
: PRIOR APPLICATION NUMBER: PCT/US99/25373
: PRIOR FILING DATE: 1999-10-08
: PRIOR APPLICATION NUMBER: GB-0004695.3
: PRIOR FILING DATE: 2000-02-28
: PRIOR APPLICATION NUMBER: PCT/US/05928
: NUMBER OF SEQ ID NOS: 107
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 2272325
: TYPE: DNA
: ORGANISM: Neisseria meningitidis
US-10-018-470-1

Alignment Scores:
Pred. No.: 609
Score: 145.50
Percent Similarity: 33.48%
Best Local Similarity: 20.49%
Query Match: 3.08%
DB: 9
Caps: 40
Length: 2272325
Matches: 191
Conservative: 121
Mismatch: 342
Indels: 278
```

```
US-09-914-168-2 (1-919) x US-10-018-470-1 (1-2272325)
OY 116 ThrProLeuSerLeuGluIleuPheAlaGlnIleuSerThyGluMetGlyIleAsnPro 135
Db 180761 TCGCTTTGGCACTTGGCGACTTCCACATCCAAAGACATCCGGGTGGAAGGCTTGCAGGCT 180702
OY 136 AsnAsp-----TYrIleProGluTYrGlnGlyGluIleProAsnSer 149
Db 180701 ACCGAGCGGAGTACCGGTATTCACATCTGCCCGCAAGTCGCGACACCTACAC--- 180645
OY 150 GluValValProProThyLeuIleuProGluIleuProGluIleuIleuArgLeuTYr 169
Db 180644 -----GACACACGCGAGTCCCATGATAAAGCTGTAC 180609
OY 170 AlaArgLeuPheAsnAspGlyValAsnLYsValPro-----ArgLeuLYsAlaLYs 186
Db 180608 GCCACCGGTTTCTTTGACGAGCAGTACGCGTCCGAAATCGCGAGCGGACACTCTCTGTAC 180549
OY 187 PheTYrGlnSerSerGlnSerGlyIleuThySerAlaIleGlySerSerHisGlnLYsThr 206
Db 180548 GTTATCGAAGCCCGACCATCGCTGCTCAACATCACCGGCGCAAAATGCTGCAAAAC 180489
OY 207 GluProTYrAlaAsnIleLYsAlaIleuGluAsp-----IleThyGlnIleuSerAla 224
Db 180488 GACGCC-----ATTAGAAACCTCGATCGTTGGGCTGGCGAGTCCCAATAC 180438
OY 225 MetAsp-----LeuAsnGlySerIleProArgLeuArgGln----- 236
Db 180437 TTTATTCAGGGGACACCTCAATACGACGACTGCGCGCTGAAAGAAATACCTCGCGGC 180378
OY 237 -----ThyAlaLeuValAlaAlaIleArgAlaValGlyTYrTYrAsp 249
Db 180377 GCGAAACTCAATATCCAAATGCAACGCAAGTAAACCAACTCGCCGCAACCGGCTGCAC 180318
OY 250 IleAspLeuSerIleIleArgAsnSerIleGlyGluValAspValIleIleHisAspLeu 269
Db 180317 ATCCGATCAGCATTTGACGAGGCAATCCGCCAAATACCCAGACATGCAATTGAA--- 180261
OY 270 GlyIleuProValTYrIleAspTYrArgAlaValGluValArgGlyGluGlyAlaAspAsp 289
Db 180260 GCGAACCAAGCTTATTCGCGACGCGCAACTGATGCGG----- 180225
OY 290 LysAlaPheThyThyValAlaAspGluValProLeuLeuIleGlyAspVal----- 306
Db 180224 -----CAATGTCCCTGACGGAAGGCGGCTTTGGACATGG 180189
OY 307 -----PheHisIleGlyLYsTYrGlyIleuThyLYsAsnLeuIleGlu 320
Db 180188 CTGACACGACCAACCATTTCAACGACGAGAAATTTGCCCAAGATATGCAAAAGTAAAC 180129
OY 321 AsnAlaSerAlaGlnHisGlyTYrPheAspGlyArgTYrLeuAspArgSerValAspVal 340
Db 180128 GACTTCTACCAAAATATACGGCTACTGATTCGATTCGTCATCGCAACGACCAAAAC 180069
OY 341 IleuProAspAsnThyAlaAspValSerLeuIleTYrAspThyGlyTYrThyGlnTYrArg 360
Db 180068 AACGAGCAAAACCAACCAACACATCAAAATCACCGTCCAGAGGGGAGCGTTTCCGT 180009
OY 361 PheAspGluValValPhePheThyIleAspProLYsTYrAsnGlnIleuThyThyAspPro 380
Db 180008 TGGGCGAAAGTC-----TCCATCGAAGCGACACCAACGAA----- 179973
OY 381 AspLYsLeuProValLYsArgGluLeuLeuGluIleuLeuThyValAsnMetGlyGlu 400
Db 179972 -----GTCCCAAGCGCAACTGGAATAAAGTGTGACCAATGAAACCGGCANA 179925
OY 401 AlaTYrAsnLeuGlnAlaValAlaArgAlaLeuSerAsnAspLeuIleAlaThyArgTYrPhe 420
Db 179924 TCGTACGAAACCGACGACATACCGCTTTGGGT----- 179889
OY 421 AsnMetValAsnThyGlnIleValPheProGluArgGluIleGlnAsnAspGlnVal 440
Db 179888 -----GAGATTCAAGACCGCATGGGC 179868
```



```

OY 697 LysThrGlnAlaProGluThrTrpGlnAspLeuPro-----ValAspPheVal 713
      |||::: ||
Db 179210 GCAGAACACCTGACCGTACACCTACACAAAGCGCCAAACACTATGCCGACTTATTC 179151
      |||::: |||
OY 714 Asn-----GlyLys-----ProSerGlnGluAlaLeuGluAlaGlyValAlaValHisLys 730
      |||::: |||
Db 179150 AAGAAATACGGCAAAACCGACGACGACGAGCTTCAAAAGCTGGCTGTACAAAGT 179091
      |||::: |||
OY 731 ThrVal-----AlaAspAsnLeuValAsnPrometArgGlyTyrArg 744
      |||::: |||
Db 179090 ACCGTGGCTGGGGGGCGCAACAAACGACAGCCGCTTATGGCCGACGCGGCTAC-- 179034
      |||::: |||
OY 745 GlnArgTyrSerLeuGluValGlySerSerGlyLeuValSerAspAlaAsnMetAlaIle 764
      |||::: |||
Db 179033 -----CTGACGGGCGGTGAACCGCCGCAATC 179010
      |||::: |||
OY 765 AlaArgAlaGlyIleSerGlyValTyrSerPheGlyAspAsnAlaTyrGlySerAsnArg 784
      |||::: |||
Db 179009 GCCCTGGCTGGCAGCAACCTGCATAC-----TACTCCGCC 178974
      |||::: |||
OY 785 AlaHisGlnMetThrGly----- 790
      |||::: |||
Db 178973 ACCCACAACCAACCTGTTCTCCCGCTGAGCAAAACCTTCACGCTGATGCTCGCGCC 178914
      |||::: |||
OY 791 -----GlyIleGlnAlaGlyTyrIleTyrSerAspAsnPheAsnHisValProTyrArg 808
      |||::: |||
Db 178913 GAAGTCGGCATTCGGGGCGGTAC-----GGCAGAACCAAGCAAAATCCCTCTCTT 178863
      |||::: |||
OY 809 LeuArgPhePheAlaGlyIleGlyAspGlnSerIleArgGlyTyrAlaHisAspSerLeuSer 828
      |||::: |||
Db 178862 GAAGACTTTCACGGCGCGCGCTGGTTCGGTCCGCGATACGAAAGCGGACGCTCGGT 178803
      |||::: |||
OY 829 ProIleSerAspLysGlyTyrLeuThrGlyGlyGluValLeuAlaValGly----- 845
      |||::: |||
Db 178802 CCG-----AAAGTCTATGACGAATACGGCAAAATACAGTACGGCGCAACAAA 178752
      |||::: |||
OY 846 ThrAlaGluTyrAsnTyrGluPheMet-----LysAsp 856
      |||::: |||
Db 178751 AAAGCCAAGTCTCCGCCGACGCTCTTCCCGATGCCCGCGGAAAGACGCGGCACCC 178692
      |||::: |||
OY 857 LeuAlaGluLeuAlaValPheGlyAspIleGlyAsnAlaTyrAsp----- 870
      |||::: |||
Db 178691 GTCCGCGCTGAGCTGTTCGCCGACGACGAGCTGTGGACGGCAAAACCTACGACGAC 178632
      |||::: |||
OY 870 ----- 870
      |||::: |||
Db 178631 AACAGCAGTTCCGCGACGCGCGGACAGGTCAAAACATTTACGCGCGCGCAATACCCAT 178572
      |||::: |||
OY 871 ---LysGlyPheThrAsnAspThrLysIleGlyAlaGlyValGlyValArgTyrPheLys 889
      |||::: |||
Db 178571 AAATCCACCTTATCCACAGCAATTCGCTATTCGCCGCGCGGCTTACSTGGCTTCG 178512
      |||::: |||
OY 890 ProValGlyGluValArgValAspValAlaThrGlyValLysGlu-----GluGlyAsn 907
      |||::: |||
Db 178511 CCTTAGGCGCCATGAATTCAGTACGCTACGCTACGCGCTGAAGAAACCGGGAAGACGAA 178452
      |||::: |||
OY 908 ProIleLysLeuHisPhePheIleGlyThrProPhe 919
      |||::: |||
Db 178451 ATCCAAACGCTTCCATTCGCACTCGGCAAGAGCTTC 178416
      |||::: |||

```

Search completed: May 6, 2003, 03:37:45
 Job time : 3534 secs

